

**PU020030 (JP6139289) ON 8044**

- (19) Patent Agency of Japan (JP)
- (12) Official report on patent publication (A)
- (11) Publication number: 6-139289
- (43) Date of publication of application: 20.05.1994
- (51) Int.Cl. G06F 15/40 G10L 3/00 G11B 27/28
- (21) Application number: 4-287705
- (22) Date of filing: 26.10.1992
- (71) Applicant: Olympus Optical Co LTD
- (72) Inventor: Okano Hideo
- (54) Title of the invention: Information reproducing device
- (57) Abstract:

Purpose: To retrieve a desired recording part in recorded voice information without operating any troublesome operation.

Constitution: In a voice registration mode, a feature extracting part 40 extracts a voice parameter from a voice signal inputted from a microphone 10 and registers it through a switch 36 which is switched and controlled by a main control part 22 in a word registering part 38. In a retrieval mode, the main control part 22 controls a driving control part 28 and operates a high speed reproducing operation. Then, the feature extracting part 40 extracts the voice parameter from a reproduction signal from a head 16 and transmits it through the switch 36 to a comparator 42.

The comparator 42 compares the voice parameter registered in the word registering part 38 with the transmitted voice pattern and transmits a detection signal to the main control part 22 at the time of detecting an equal part. The main control part 22 inputs the detection signal, then transmits a stop signal to the driving control part 28 and stops the high speed reproducing operation.

### **[Claims]**

#### **[Claim 1]**

An information reproducing device including a characteristic standard pattern registration means to register preliminary a characteristic standard pattern of speech information of predetermined length used as search information, a characteristic pattern extraction means to analyze a predetermined audio signal read from a recording medium with which speech information longer than speech information of the mentioned above predetermined length is recorded and to extract a characteristic pattern, a judging means that compares an extracted characteristic pattern with a characteristic standard pattern of speech information of predetermined length registered preliminary and judges the similarity.

#### **[Claim 2]**

The information reproducing device according to claim 1 characterized by extracting a characteristic pattern by coding a digitized audio signal and changing this coded data.

[Claim 3]

The information reproducing device according to claim 1 characterized by comparing with a characteristic standard pattern of speech information of predetermined length an analog voice signal by which fast reproduction was carried out, changing a frequency axis, so that the similarity can be judged and recognizing and searching a fast reproduction signal.

**[Detailed description of the invention]**

[0001]

[Industrial application] This invention relates to the information reproducing device used for the device that accumulates recording and reproduces the sound sent by the dictaphone device, a telephone line, radio, etc. that record and reproduce a sound.

[0002]

[Description of the prior art] Recently, what is in a manager's status especially with the clerical work rationalization in an office must manage many workloads and almost all time is spent on this. For this reason, if it is in creation of a letter, a report while it cannot perform writing a manuscript by hand each time, but the contents of a manuscript are preliminary recorded on a tape instead of this and a secretary plays this, a type and a thing that carries out writing copy are performed.

[0003] In such case, a note device that is indicated by JP 60-25039 A and JP 63-112886 A is one of those are used, for example. It is thought that this note device is useful to perform type operation while a secretary hears

the contents of a tape. For this reason, many functions to cross variously are demanded.

[0004] For example, in such a note device, there is a demand of a witness distinguishing form of a message, existence of a priority, many cautions, etc. that were recorded preliminary to what it is going to tell to a copyist and making him like to tell it. While the number, the position and various forms of the message that exists on a recording medium are discriminable as a feature further demanded in a note device, there is a thing of enabling it to search for a desired message easily on a recording medium.

[0005] Then, the witness was made to record a tone or a cue signal as a recognition signal conventionally, for example after a series of texts later mentioned using dictaphones, such as a tape recorder, in order to distinguish from other texts that follow. When a related message needed to be recorded, he was trying to record the different tone or cue signal from the mentioned above recognition signal in the case of the record.

[0006] The typical operation method for recording a recognition signal, when a message to inform a copyist on the way while the witness is performing record by oral statement etc. occur, when a cue signal is recorded as a signal that pushes and identifies cue and the end of oral statement message (a series of texts) record is shown, the mentioned above recognition signal pushes another cue and records a different recognition signal.

[0007] Thus, before a copyist writes down, he can perform fast reproduction of information and can hear a witness's attentions preliminary. Since the relative location of a recording medium and a cue signal are displayed on the display part of a note device at the time of this fast reproduction, the copyist can distinguish information to hear visually.

[0008] The system configuration of such a note device takes composition as shown, for example on drawing 7 (B). This drawing simplifies drawing 1 of mentioned above JP 60-25039 A.

[0009] That is, in sound recording operation, the audio signal inputted from the microphone 10 is amplified with the microphone amplifier 12 and is recorded on recording medium that is not represented by the head 16 via the sound recording playback amplifier 14. In reproduction motion, the signal detected from the recording medium by the head 16 is amplified with the sound recording playback amplifier 14, is amplified with the power amplifier 18 and is sent to the loudspeaker 20. Control of operation of these series is performed by the control part 22. This control part 22 detects the state of the switch of sound recording and playback by the interface 24 and sends a control signal to the display part 26 or the driving control part 28 again. The driving control part 28 controls the motor 30.

[0010] When a cue button is pushed at the time of sound recording, a signal is sent to the control part 22 from the interface 24 and the control part 22 sends a cue control signal to the bias oscillating circuit 32 according to this

and makes a cue signal record on a recording medium by the head 16. A cue signal is detected from the cue signal detecting part 34 by the head 16 and the control part 22 makes predetermined operation that sent and the mentioned above control signal in the display part 26 or the driving control part 28 perform according to it at the time of reproduction or fast reproduction.

[0011]

[Problems to be solved by the invention] However, in the conventional note device, when looking for the desired part in the text by which oral statement record was carried out, the witness has to perform reproduction or double-speed reproduction and has to perform operation of actually hearing it with an ear and finding out the mentioned above desired part for the contents with slight accuracy. Generally, when such, by the time he finds a desired part, complicated operations as repeating operation of reproduction, a fast-forwarding and rewinding will be needed.

[0012] This invention was made in view of the mentioned above point and an object of this invention is to provide the information reproducing device that can search a desired part, without performing complicated operation of repeating operation of reproduction, a fast-forwarding and rewinding, when finding out the desired part in the recorded speech information.

[0013]

[Means for solving the problem] In order to achieve the mentioned above purpose, an information reproducing device according to this invention, a characteristic

standard pattern registration means to register preliminary a characteristic standard pattern of speech information of predetermined length used as search information, a characteristic pattern extraction means to analyze a predetermined audio signal read from a recording medium with which speech information longer than speech information of the mentioned above predetermined length is recorded and to extract a characteristic pattern, it includes a judging means that compares an extracted characteristic pattern with a characteristic standard pattern of speech information of predetermined length registered preliminary and judges the similarity.

[0014]

[Function] Namely, in order to look for the part of a request of the speech information currently recorded on the recording medium according to the information reproducing device of this invention, registering preliminary the characteristic standard pattern of the speech information of the predetermined length used as search information, analyzing the predetermined audio signal read from the recording medium and a characteristic pattern is extracted, the extracted characteristic pattern is compared with the characteristic standard pattern of the speech information of the predetermined length registered preliminary and the similarity is judged.

[0015]

[Example] Next, the example of this invention is described.

(The 1st example) Drawing 1 is a drawing showing the composition of the 1st example of this invention and the same reference numbers as on drawing 7 (B) are attached to the same parts in the drawing.

[0016] That is, in sound recording operation, after the audio signal inputted from the microphone 10 is amplified with the microphone amplifier 12, it is recorded on an recording medium that is not represented by the head 16 through the sound recording playback amplifier 14. In reproduction motion, after the signal detected from the mentioned above recording medium that is not represented by the mentioned above head 16 is amplified with the mentioned above sound recording playback amplifier 14, it is further amplified by the power amplifier 18 and is sent to the loudspeaker 20. These series of operations are controlled by the main control part 22 by the interface 24, this main control part 22 detects states, such as recording switch REC, regeneration switch PLAY, the cue buttons CUE1 and CUE2, the voice registration button RRREC and retrieval button SV and sends a control signal to the display part 26 or the driving control part 28 further again. According to the mentioned above control signal from the main control part 22, the display part 26 performs the display variously, such as a cue signal display and a counter display and the driving control part 28 performs drive controlling of the motor 30.



[0017] On the other hand, if the mentioned above cue button CUE is pushed at the time of the mentioned above sound recording, a signal will be sent to the mentioned above main control part 22 from the mentioned above interface 24. The main control part 22 sends a cue control signal to the bias oscillating circuit 32 according to this and makes a cue signal record on the recording medium that is not represented by the mentioned above head 16. and a cue signal is detected by the cue signal detecting part 34 by the mentioned above head 16 and the mentioned above main control part 22 sends a control signal to the display part 26 or the driving control part 28 and makes predetermined operation perform according to the detection at the time of reproduction or fast reproduction. The above is the same as that of the conventional device. The information reproducing device of this example has the still more nearly following composition.

[0018] That is, depression of the voice registration button RRREC will send a signal to the mentioned above main control part 22 from the mentioned above interface 24. The main control part 22 switches the switch 36 to the word registering part 38 side according to this. And it is inputted into the audio feature extracting part 40 by the microphone amplifier 12, the vocal parameter extracted by this feature extracting part 40 is inputted into the mentioned above word registering part 38 by the mentioned above switch 36 and the audio signal inputted from the mentioned above microphone 10 is registered there.

[0019] Next, if retrieval button SV is pushed, a signal is sent from the mentioned above interface 24 and the mentioned above main control part 22 will send a fast reproduction actuating signal to the driving control part 28 and will perform fast reproduction operation. After detecting a reproducing signal from the magnetic tape that is not represented by the mentioned above head 16 and amplifying by the mentioned above sound recording playback amplifier 14 at this time, a vocal parameter is outputted by the mentioned above feature extracting part 40. Here, when the mentioned above main control part 22 carries out switching control of the mentioned above switch 36, the vocal parameter is sent to the comparator 42. If this comparator 42 compares the vocal parameter registered into the mentioned above word registering part 38 with the voice pattern sent from this feature extracting part 40 and a part in agreement is detectable, it will send a detecting signal to the mentioned above main control part 22. If this detecting signal is inputted, the main control part 22 will send a stop signal to the mentioned above driving control part 28 and will stop fast reproduction operation.

[0020] When finding the part that a witness wants to rerecord, performing double-speed reproduction or reverse fast reproduction as mentioned above, by finding out the desired place by audio registration, without performing complicated operation of repeating operation of playback, a fast-forwarding and rewinding, a desired voice recording part can be searched and addition of a cue mark and recorded audio edit can be easily performed now.

[0021]By the way, magnetic tape is used as the mentioned above recording medium that is not represented and when searching performing fast reproduction, an audio signal will change by fast reproduction. For this reason, since the signal that carried out fast reproduction and the registered sound become a completely different audio signal, it is necessary to incorporate a certain technique that can be searched also with the signal by which fast reproduction was carried out. Then, when carrying out sound pattern matching to sound that was registered and fast reproduction was carried out, it is made to coincide a time-axis in this example by changing the sampling frequency at the time of voice registration and the sampling frequency of fast reproduction. At the registration time, it shall be 8 kHz and, specifically, a sampling frequency shall be 16 kHz at the time of double-speed reproduction. The concrete composition of the 1st example that incorporated such a technique is shown on drawing 2 and drawing 3.

[0022] That is, the microphone amplifier 12 was connected to the microphone 10 and the magnetic head 16 was connected to this microphone amplifier 12 by the sound recording playback amplifier 14 and the loudspeaker 20 is connected to the sound recording playback amplifier 14 by the power amplifier 18. The reference number 32 is a bias oscillating circuit and 44 is an external microphone connection terminal.

[0023] For example, it is connected to the sound recording playback amplifier 14 or the power amplifier 18 and the main control part 22 that includes a microcomputer performs control of sound recording and playback. The remote control operating parts 48, such as the keyboard 46, a foot switch and a hand controller are connected between each terminals D0-D3 of output port PD of this main control part 22 and each terminals K0-K3 of input port PK. Here, the mentioned above keyboard 46 has the recording key REC1, the rewinding key REW1, the reproduction key PLAY1, the fast forwarding key FF1, the voice registration key RRREC, the cue key CUE, the erase key ERASE, the stop key STOP and the search key SV. The remote control operating part 48 has the recording key REC2, the rewinding key REW2, the reproduction key PLAY2, PLAY3.

[0024] The display circuit 26 is connected to each terminal A0 of input/output port PA of the mentioned above main control part 22 - A3 and terminal B-2 of the input/output port PB and B3. This display circuit 26 is constituted so that status displays, such as sound recording, playback, rewinding and fast-forwarding, can be performed by LED.

[0025] The solenoids 52A, 52B are connected to each terminals E0-E3 of output port PE of the mentioned above main control part 22 by the solenoid driving circuit 50 that accomplishes a part of mentioned above driving control part 28. It is for these solenoids 52A, 52B driving a mechanism part in each mode and the solenoid 52A is driven at the time of sound recording

reproduction mode and drives the solenoid 52B at the time of fast forwarding rewinding mode. The motor 30 is connected to the terminals F0-F3 of output port PF of the main control part 22 by the motor drive circuit 54 that accomplishes a part of mentioned above driving control part 28.

[0026] The backspace set part 56 is connected to the terminal B0 of the input/output port PB of the mentioned above main control part 22. This backspace set part 56 is for setting up time to rewind only predetermined time, when it halts by reproduction mode and is again considered as reproduction mode.

[0027] Each terminal G0 - G3 of output port PG of the mentioned above main control part 22 connected with the bias oscillating circuit 32 and the terminal G0 generates a CUE output signal according to cue key CUE operation of the mentioned above keyboard 46 at the time of sound recording mode. The terminal G1 generates an alarm signal, an AC output. It connects with the sound recording playback amplifier 14 and the terminal G2 generates the signal that delays the standup of the output of the sound recording playback amplifier 14 predetermined time at the time of sound recording mode. Terminal G3 is connected to the bias oscillating circuit 32. According to operation of recording switch REC1 of the mentioned above keyboard 46 and REC2, the REC output signal of sound recording starting is generated.

[0028] The REEL input for pre-end detection is given to the input port H0 of the mentioned above main control part 22. In this case the REEL input pulse form rotation signal of the motor according to rotation of the reel that is not represented is used. While connecting the change-over switch 58 and connecting the mentioned above bias oscillating circuit 32 as a source of a cue signal to this switch 58, the mentioned above cue signal detecting part 34 is connected to input port INT. This cue signal detector circuit 34 detects a cue signal and is trying to input this detect output into the mentioned above input port INT from the reproducing output of the magnetic head 16 at the time of reproduction mode. And the oscillating circuit 60 is connected to terminal OSC0, OSC1 of the mentioned above main control part 22.

[0029] Terminal SR of the digital signal processing part (DSP) 62 is connected to the mentioned above switch 58 again. On the other hand, AD / DA converter 64 is connected at the nodal point of the mentioned above microphone amplifier 12 and the sound recording playback amplifier 14. And between both FSXs and a FSR terminal is connected between the terminal DR, the terminal POUT, terminal DX and terminal PIN and both the CLK terminal, respectively and this AD / DA converter 64 and the mentioned above digital processing part 62 transmit and receive the digitized sound data. The oscillating circuit 66 that generates the clock that determines access speed is connected to both the CLK terminal.

[0030] This oscillating circuit 66 is connected with the terminal H3 of the mentioned above main control part 22. By the signal generated from this terminal H3, a clock frequency can be changed in the internal frequency divider 66A. Also, this oscillating circuit 66 is connected to FS terminal of mentioned above AD / DA converter 64 and the timing control circuit 68. It is connected to the terminals FSX and FSR of the mentioned above digital processing part 62 and the AD/DA converter 64 and this timing control circuit 68 is controlling transmission timing.

[0031] As for the mentioned above digital signal processing part 62, the address terminal and the data terminal are connected to the memory 70. The decoder 72 connected to the control signal of this digital signal processing part 62 generates a selection signal and it is connected, so that it may be inputted into terminal CS of the mentioned above memory 70.

[0032] The data terminal of the mentioned above digital signal processing part 62 is connected to the buffer 74 again and this buffer 74 is connected with the mentioned above decoder 72 too. The terminals D1-D8 of this buffer 74 are connected to the terminals C0-C7 of input/output port PC of the mentioned above main control part 22. This buffer 74 is used in order to enable transmission and reception of data between the mentioned above main control part 22 and the mentioned above digital signal processing part 62 that differ in a process speed. Next, the operation in such composition is explained with reference to the schematic diagram of processing on drawing 4.

[0033] It passes along the course A on drawing 4 by making the characteristic standard pattern of the sound used as introduction and search information into the voice registration mode registered preliminary. The flow of processing in this voice registration mode becomes like the flow chart shown on drawing 5 (A).

[0034] First, if the keyboard 46 is voice registration switched on RRREC, the main control part 22 will detect that from PD port and PK port and will input a sampling switching control signal into FT terminal of the oscillating circuit 66 from the output port H3. With this signal, the oscillating circuit 66 sets a sampling frequency as 8 kHz by the internal dividing switch circuit 66A (Step S11).

[0035] And after the audible signal inputted from the microphone 10 is amplified with the microphone amplifier 12, it is inputted into AD / DA converter 64. This AD / DA converter 64 sample an input analog audible signal by 8 kHz of sampling frequencies and changes it into a digital signal. This digital audible signal is sent to the pretreatment part 76 inside the digital signal processing part 62. In this pretreatment part 76, the inputted sound data in a high emphasis filter (pre-emphasis) after performing noise rejection using an adaptation filter, a VOX signal will be outputted, if it judges from spectrum, a level, whether the audible signal inputted from the microphone 10 is a sound and judges that it is silent (VOX) (Step S12). If this VOX signal is no longer detected (Step S13), sound data will be gathered in the block data for every 10 msec - 30



msec and will be sent to the feature parameter extracting part 40.

[0036] Processing of this feature parameter extracting part 40 inside is shown on drawing 6 (A). That is, the block sound data sent from the pretreatment part 76 is inputted into the acoustic analysis part 40A. This acoustic analysis part 40A calculates an autocorrelation function from linear predictive coding. At this time, linear predictor coefficients  $\{a_1, a_2, a_3, \dots, a_p\}$  and autocorrelation function  $\{r(1), r(2), r(3), \dots, r(p)\}$  can be found using reverse spectrum coefficient calculation part 40B, block data of every feature parameter reverse spectrum coefficient  $\{A_1, A_2, A_3, \dots, A_p\}$  accumulating (Step S14). This reverse spectrum coefficient is accumulated in the word registering part 38 as retrieved data until that storage capacity fills (several seconds) (Step S15). And if a storage capacity fills, registration of the voice input signal from the microphone 10 will be finished namely, disregarded (Step S16). Next, explanation is added about a linear predictive method (linear predictive coding).

[0037] The thing of analysis by a linear predictive method (linear prediction method) is usually called LPC analyzing. LPC is linear predictor coefficients (linear prediction coefficients) and the information about a spectrum is collected by the coefficient that is about ten pieces. The feature of this method searches for the generation process of a sound until the sound that becomes the origin of the voice uttered in the sound source comes out from a mouth through a vocal tract from a voice waveform.

It is the most effective means also as a synthesizing voice method not only as an analysis method of a spectrum.

[0038] The voice waveform that reads  $N$  pieces (150 pieces - about 300 pieces) for every (there are usually about 100 microsecond - 125 microsecond)  $\Delta t$  second is obtained in the continuous voice waveform  $s(t)$  as the digital data  $\{s(\Delta t), s(2\Delta t), \dots, s(N\Delta t)\}$ . The thing of operation that reads such a waveform is called a sampling and the thing of read interval  $\Delta t$  is called a sampling period. In order to avoid complication,  $\Delta t$  is omitted and the data point in time  $t=n\Delta t$  is written to be  $s(n)$ . Only  $p$  pieces are as follows using the past value than using it in the value  $s(n)$  in a certain time  $n$  (time is  $n\Delta t$  actually), it predicts (here,  $s'(n)$  is considered as a predicted value to  $s(n)$ ).

$$s'(n) = -a_1s(n-1) - a_2s(n-2) - \dots - a_ps(n-p)$$

[0039] Since this formula expresses  $s(n)$  with the linear combination of the past value rather than it, it is called linear prediction. These linear predictor coefficients decide that predicted value  $s'(n)$  becomes a value near  $s(n)$  as much as possible. The least-squares method is generally used as the effective method.

That is, if a prediction error is set to  $e(n)$  now, it will be

$$e(n) = s(n) - s'(n)$$

$$= s(n) + a_1s(n-1) + a_2s(n-2) \dots + a_ps(n-p)$$

[0040]  $e^2(n)$  about total time  $n$  of  $y^2 = e^2(1) + e^2(2) + e^2(3) + \dots$  a coefficient is decided that becomes the minimum. If the condition is searched, linear predictor coefficients can be found as follows.

[0041]

[Equation 1]

$$\left. \begin{aligned} r(0) a_1 + r(1) a_2 + r(2) a_3 + \dots + r(p-1) a_p &= -r(1) \\ r(1) a_1 + r(0) a_2 + r(1) a_3 + \dots + r(p-2) a_p &= -r(2) \\ \vdots & \\ r(p-1) a_1 + r(p-2) a_2 + r(p-3) a_3 + \dots + r(0) a_p &= -r(p) \end{aligned} \right\}$$

This is called a normal equation  $r(0), r(1), \dots, r(p)$  are called the autocorrelation function of signal. This autocorrelation function,

[0042]

[Equation 2]

$$r(m) = \sum_{n=1}^{N-m} s(n) s(n+m) \quad m = 0, 1, 2, \dots, p$$

[0043] However, this adds  $\{s(n)\}$  itself and the things to which only  $m$  sample shifted the whole for it. At the time of  $m=0$ , since  $r(0)$  is the double sum of a signal, it is the energy of  $s(n)$ . Thus, a sampled value  $\{s(1), s(2), \dots, s(N)\}$  the simultaneous equations that calculate the autocorrelation function of an upper equation from, then are made are solved and linear predictor coefficients are calculated.

[0044] A reverse spectrum coefficient becomes settled from the formula of  $s^2$  of total of this autocorrelation function  $\{r(0), r(1), \dots, r(p)\}$ , linear predictor coefficients  $\{a_1, a_2, \dots, a_p\}$  and prediction error to the next.

$$s^2 = r(0) A_0 + 2r(1) A_1 + 2r(2) A_2 + \dots + 2r(p) A_p$$

[0045] Next, it is set in the voice search mode that searches the registered sound. This voice search mode takes the course B in drawing 4. The flow of processing of this voice search mode becomes like the flow chart shown on drawing 5 (B).

[0046] First, if search key SV of the keyboard 46 is turned on, the main control part 22 will detect it from PD port and PK port and will input a sampling switching control signal into FT terminal of the oscillating circuit 66 from the output port H3. Thus, it is set as the sampling frequency corresponding to fast reproduction. In this example, if it is search by double-speed reproduction, it will be set as 16 kHz (Step S21) (set to 24 kHz in 3X). The main control part 22 outputs a fast reproduction actuating signal to PE port and PF port, for example, is made to be reproduced by double speed (Step S22).

[0047] And after the audible signal inputted from the head 16 is amplified with the sound recording playback amplifier 14, it is inputted into AD / DA converter 64. AD / DA converter 64 samples by 16 kHz of sampling frequencies and changes an analog audible signal into a digital signal. This digital audible signal is sent to the pretreatment part 76 inside the digital signal processing part 62. By this pretreatment part 76, after the sound data part is summarized to the block data for every 10 msec - 30 msec and a high emphasis filter (pre-emphasis) lets it pass and it performs noise rejection using an adaptation filter etc., it is sent to the feature-

parameter extracting part 40. In this feature-parameter extracting part 40, like the time of voice registration mode, the block sound data signal  $Z(n)$  sent from this pretreatment part 76 is inputted into the acoustic analysis part 40A and is calculated in an autocorrelation function from linear predictive coding. At this time, linear predictor coefficients  $\{a_1, a_2, a_3, \dots, a_p\}$  and autocorrelation function  $\{r(1), r(2), r(3), \dots, r(p)\}$ , it can be found next reverse spectrum coefficient calculation part 40B using reverse spectrum coefficient  $\{A_1, A_2, A_3, \dots, A_p\}$  asking (Step S23), shortly, it is sent to the pattern matching part 42 as a comparator.

[0048] In the pattern matching part 42, in order to compare the pattern of a standard voice and the voice pattern to search and to evaluate the form of the spectrum of patterns, spectrum distance is evaluated. As this interval scale, there are an LPC cepstrum interval scale, a weighting likelihood ratio (WLR) measure maximum spectrum distance scale method, etc. In this example, square sum  $s^2$  whose playback voice pattern is each prediction residual (prediction error) as the standard voice pattern of the word registering part 38 is calculated, from this prediction residual, the value drawn by likelihood  $l = -\log s^2$  is used as similarity and pattern matching is performed in quest of the spectrum distance of a playback voice pattern and a standard voice pattern. Drawing 6 (B) is a drawing showing processing of this pattern matching. Here, input pattern A is expressed with  $\{a_1, a_2, a_3, \dots, a_l\}$  to 10 msec - 30 msec as a feature vector of the block data by which the part division was carried out and the standard voice

pattern B of the word registering part 38 expresses with  $\{b_1, b_2, b_3, \dots, b_J\}$ . The time of considering the flat surface that consists of these patterns A and B, the series of lattice point  $C = (i, j)$  on this flat surface.

$F = C_1, C_2, C_3, \dots, C_k, \dots, C_K \quad C_k = (i_k, j_k)$

[0049] It becomes. Spectrum distance of  $a_i$  and  $b_j$  is set to  $d(c) = d(i, j)$ . In pattern matching, although it searches for the minimum of this spectrum distance, pattern matching is changed into the problem of dynamic programming and finds the cumulative distance  $g$  here for simplification.

[0050]

[Equation 3]

$$g(C_k) = g(i, j) \\ = \begin{bmatrix} g(i-2, j-1) + 2d(i-1, j) + d(i, j) \\ g(i-1, j-1) + 2d(i, j) \\ g(i-1, j-2) + 2d(i, j-1) + d(i, j) \end{bmatrix}$$

calculation is performed repeatedly.

[0051] Thus, when carrying out pattern matching and fulfilling the conditions that distance is certain, shifting a standard pattern and a playback voice by one frame (block data unit of 10 msec – 30 msec), it judges that the corresponding word existed.

[0052] That is, if it returns to the mentioned above step S23 and there is a pattern in agreement after clearing an input pattern (Step S25), if there is no standard pattern that is in agreement with an input pattern (Step S24), a code area and inclination restrictions as shown on drawing 7 (A) are provided and matching is judged

(Step S26). And if it has separated from this code area and inclination restrictions, it progresses to the mentioned above step S25 and that pattern will be asked for the accumulation g, namely, cumulative distance, if it has not separated (Step S27). And if the cumulative distance g has not reached a retrieval window (Step S28), when it is judged that it returned to the mentioned above step S23 and the retrieval window was reached. A detecting signal is sent to the INT terminal of the main control part 22 from the pattern matching part 42 and so, the main control part 22 sends a control signal to the driving control part 28 from PE port and PF port in order to perform a reproduction stop (Step S29). During search, the control signal for making a display that makes the display part 26 blink a repeat display from PA port of the main control part 22 accomplish is sent.

[0053] As mentioned above, when finding out the part that a witness wants to record when recording and adding a voice message and a voice message recognition signal later, the sound equivalent to the sound currently recorded on the desired place is registered and the desired place can be found and searched with performing speech recognition in the tape position by audio record, performing double-speed reproduction or reverse fast reproduction. When what put in the cue signal specially until now and was performing a message and directions to the secretary inputs a sound from a microphone, for example during “addressing to a secretary” and sound recording, it enables a desired part to search the inputted sound to reliance.

[0054] The above description explained the technique of changing the sampling frequency at the time of voice registration and fast reproduction, so that it could refer also to the signal by which fast reproduction was carried out, but it is also possible to use the following techniques, for example.

(1) Perform a word registration sound or the interval conversion process (frequency conversion processing) of a reproducing signal before electrical parameter extraction.

(2) Perform frequency-axis conversion data processing at the time of pattern matching.

(3) In order that a gain may go up at the time of fast reproduction, perform gain correction.

(2nd example) Next, the case where this invention is applied to digital sound recording reproducing device is explained.

[0055] Namely, after the 2nd example carries out digital sound signal conversion, by digital signal processing, it codes a sound, extracts an audio feature parameter and stores it to a tape, a disk, semiconductor memory, etc. And voice search changes the parameter that can be found at the time of the mentioned above coding, performs audio pattern matching and searches a sound.

[0056] In drawing 8, the audio signal first inputted from the microphone 80 is inputted into the 1st amplifier 82. The output terminal of this 1st amplifier 82 is connected to the terminal ANALOGIN of the AD/DA converter 84.



The terminal ANALOGOUT of the AD/DA converter 84 and the input terminal of the 2nd amplifier 86 are connected and the loudspeaker 88 is connected to the output terminal of this 2nd amplifier 86. That is, after changing the audio signal digitized into an analog signal by the AD/DA converter 84, sending it to the 2nd amplifier 86 and amplifying it with this amplifier 86, it is outputted to the loudspeaker 88.

[0057] Between the AD/DA converter 84 and the digital signal processing part 90, terminal DT, the terminal PCMIN, the terminal PCMOUT, the terminal DR and the terminal CLK and the terminal SCLK, the terminal FSR, terminal FSX and the terminal RFS and the terminal TFS are connected, respectively, the signal by which digital conversion was carried out by the AD/DA converter 84 is transmitted to the digital signal processing part 90 or the signal processed by the digital signal processing part 90 is transmitted to the AD/DA converter 84.

[0058] The terminal ADDRESS of the mentioned above digital signal processing part 90 is connected to the terminal ADDI of the 1st buffer 92 and the terminal ADDI of the 2nd buffer 94. Terminal DATA of the digital signal processing part 90 is connected to the terminal DATAI of the 1st buffer 92 and the terminal DATAI of the 2nd buffer 94, each terminals RD, WR, DMS of the digital signal processing part 90 are connected to the terminals RD, WR, BMS of the 1st buffer 92 and the terminals RD, WR, BMS of the 2nd buffer 94.

[0059] The terminal ADDO of the 1st buffer 92 and the terminal ADDO of the 3rd buffer 96 and the ADDRESS terminal of the 1st memory 98 are connected and the terminal DATAO of the 1st buffer 92 and the terminal DATAO of the 3rd buffer 96 and the data terminal of the 1st memory 98 are connected. Similarly, the terminal ADDO of the 2nd buffer 94 and the terminal ADDO of the 4th buffer 100 and the ADDRESS terminal of the 2nd memory 102 are connected and the terminal DATAO of the 2nd buffer 94 and the terminal DATAO of the 4th buffer 100 and the data terminal of the 2nd memory 102 are connected. Each CS terminal of the 1st buffer 92 of the above and the 3rd buffer 96 and CS terminal of the 1st memory 98 are connected and CS terminal of the 2nd buffer 94 and the 4th buffer 100 and CS terminal of the 2nd memory 102 are connected.

[0060] The terminal ADRB of the main control part 104 is connected to the terminal ADDI of the 3rd buffer 96 and the terminal ADDI of the 4th buffer 100. Terminal DAB of the main control part 104 is connected to the terminal DATAI of the 3rd buffer 96 and the terminal DATAI of the 4th buffer 100. Each of the main control part 104 terminals IR and IW are connected to the terminals RD and WR of the 3rd buffer 96 and the terminals RD and WR of the 4th buffer 100.

[0061] Terminal SB1 and terminal SB2 of the main control part 104 are connected to the terminal A and the terminal B of the decoder 106. The terminal Y0 of this decoder 106 is connected to terminal CE of the 1st buffer 92.

Similarly, connect the terminal Y1 of the decoder 106 to terminal CE of the 2nd buffer 94, the terminal Y2 is connected to terminal CE of the 3rd buffer 96 and the terminal Y3 is connected to terminal CE of the 4th buffer 100, respectively.

[0062] The IOR terminal of the mentioned above digital signal processing part 90 is connected to the INT terminal of the main control part 104. It is connected to the display for indication 108 and PA port terminals of the main control part 104 display each operating state. PB port terminals of the main control part 104 are connected with the keyboard 110. And terminal IRQ, terminal INT and terminal SIO, terminal SIO, the terminal SCL and the terminal CLK are connected between the main control part 104 and the memory cartridge 112.

[0063] Terminal BF of the 1st buffer 92 and the 2nd buffer 94 is connected to the logic circuit 114. And this logic circuit 114 output terminal and the terminal FB of the main control part 104 are connected. Next, operation of the digital sound recording reproducing device of such composition is explained.

[0064] In introduction and sound recording mode, if the REC switch of the keyboard 110 is pushed, the main control part 104 will detect a recording start through PB port and will transmit the instruction code that makes compression processing preparations to the digital signal processing part 90 by the 3rd buffer 96 and the 1st buffer 92.

[0065] Here, after the audio signal inputted from the microphone 60 is amplified with the 1st amplifier 61, it is sent to the AD/DA converter 84 and analog-to-digital conversion is carried out by this AD/DA converter 84. The signal by which digital conversion was carried out by this AD/DA converter 84 is transmitted to the digital signal processing part 90. In the digital signal processing part 90, after packing the digitized audio signal into the block data of 10 msec – 30 msec, an audio feature parameter is extracted from this block data using an analyzed type voice encoding system like a CELP coding mode and a parameter is quantized.

[0066] The main control part 104 sends a signal to CE terminal of the 1st buffer 92 preliminary by sending a control signal to the decoder 106 from the terminals SB1 and SB2 and this 1st buffer 92 enables it to receive data on the other hand.

[0067] And the digital signal processing part 90 transmits the data coded by the data terminal to the 1st buffer 92. The 1st buffer 92 sends the signal with transmitted data to a memory from terminal BF to EB terminal of the main control part 104, after transmitting this data to the 1st memory 98.

[0068] According to this, by sending the signal of terminal SB1 and SB2, the main control part 104 sends a control signal to CE terminal of the 2nd buffer 94 and the 4th buffer 100 from the terminal Y1 of the decoder 106 and Y3 and changes it into the state in which a data input/output is possible.

[0069] And the main control part 104 receives the coding data of the 1st memory 98 by the 3rd buffer 96 and this is transmitted to the memory cartridge 112 by the serial terminals IRQ, SIO and SCL.

[0070] While the main control part 104 is performing the mentioned above work, the digital signal processing part 90 transmits the data that calculated the audio signal block of the following input signal and was coded to the 2nd buffer 94.

[0071] If coding data is transmitted to the 2nd memory 102 from the 2nd buffer 94, a control signal will be outputted to the logic circuit 114 from BF terminal and it will inform that data transfer was carried out to EB terminal of the main control part 104 from the output terminal of this logic circuit 114.

[0072] According to this, by sending the signal of terminal SB1 and SB2, the main control part 104 sends a control signal to CE terminal of the 1st buffer 92 and the 3rd buffer 96 from the terminal Y0 of the decoder 106 and Y2 and changes it into the state in which a data input/output is possible.

[0073] After that, the STOP button of the keyboard 110 is pushed and the mentioned above operation is repeated similarly and performed until the main control part 104 recognizes it and sends a stop signal to the digital signal processing part 90 through a buffer. That is, the data transfer between the digital signal processing part 90 and the main control part 104 passes 2 courses of the course by the 1st buffer 92, the 1st memory 98 and the 3rd buffer 96 and the course by the 2nd buffer 94, the

2nd memory 102 and the 4th buffer 100 by turns and is performed.

[0074] Thus, since linear predictor coefficients are accumulated as data, the thing using analysis coding as CELP whose coded data can be changed into the feature parameter of the sound for pattern matching. It has the feature at the place that can be searched at high speed as compared with the tape recorder of the 1st example that also mentioned above voice search.

[0075] It is a functional block diagram of the digital sound recording reproducing device of this example of drawing 9. The flow shown in the course A in the drawing is in voice registration mode and the course B shows the flow of voice search mode.

[0076] In voice recording mode, after the audio signal amplified with the microphone amplifier that inputs a sound and is equivalent to the above 1st amplifier 82 from the microphone 80 first is changed into a digital signal with A/D converter 84A constituted in the AD/DA converter 84, it is sent to the digital signal processing part 90. This digital signal processing part 90 takes composition as shown on the drawing, first, a signal is sent to the coding part 116 and the analysis coding of the above is carried out. And this coded data is sent to the main control part 104 by the temporary storage part that is equivalent to the mentioned above memory 98, 102 with the switches S1-S3 that includes the mentioned above buffer and a decoder and is recorded on the recording medium equivalent to the mentioned above memory cartridge 112.

And it is longer than the speech information registered into the word registering part 122, namely, much speech information is usually recorded on this recording medium.

[0077] In sound reproduction mode, by the main control part 104, the switches S1-S3 and the temporary storage part 98, 102, the coding data from the recording medium 112 is sent to the decoding part 118 of the digital signal processing part 90 and is decoded. And it is changed into an analog signal with D/A converter 84B constituted in the decoded audio signal AD/DA converter 84 and after being amplified with the power amplifier equivalent to the above 2nd amplifier 86, it is outputted to the loudspeaker 88.

[0078] In voice registration mode, after the audio signal amplified with the microphone amplifier that inputs a sound and is equivalent to the above 1st amplifier 82 from the microphone 80 is changed into a digital signal with A/D converter 84A constituted in the AD/DA converter 84, it is sent to the digital signal processing part 90. In this digital signal processing part 90, first, a signal is sent to the coding part 116 and the above analysis coding is carried out. This coded data is sent to the feature-parameter converter 120, it changes into a feature parameter using the linear predictor coefficients included in data, the same reverse spectrum coefficient as the above 1st example is obtained and it registers with the word registering part 122 for several seconds. At this time, registration is performed by voice starting like the above 1st example.

[0079] The coding data from the recording medium (memory cartridge) 112 with which the recording place was managed from the main control part 104 in voice search mode by the switches S1-S3 and the temporary storage parts 122A, 112B. It is transmitted to the feature-parameter converter 120, is changed into a feature parameter like the above and is sent to the pattern matching part 124. The method of this pattern matching takes the same method as an example as shown in the above 1st example. Predetermined stop processing is performed pointing out that a detecting signal will be sent to the INT terminal of the main control part 104 and a recording place will be found, if pattern matching is detected.

[0080] Since the parameter of linear prediction is included in the coded data in digital sound recording reproducing device, this is used, it is possible to change into the feature parameter of the sound for pattern matching from linear predictor coefficients and voice search also has that there are few operation amounts and an effect by random access that can be searched at high speed as compared with the above tape recorder of the 1st example.

[0081] If the view of this invention is promoted, application of deleting the specific language by which speech recognition was carried out or replacing by another language is possible too. For example, deletion of the language corresponding to the registered language can be automatically performed by registering language to delete and pushing a deletion button during reproduction of the portion in which the language is



included. Thus, it is dramatically effective when there is much language to delete.

[0082]

[Effect of the invention] As explained in full details above, according to the information reproducing device of this invention, finding the requested part of the recorded speech information because the desired place can be found by audio registration, the part of the mentioned above request can be searched without performing troublesome operation of repeating operation of reproduction, fast-forwarding and rewinding.

### **[Brief description of the drawings]**

[Drawing 1] is a block diagram of the information reproducing device of the 1st example.

[Drawing 2] is a drawing showing the half of a circuitry drawing where the information reproducing device of the 1st example is detailed.

[Drawing 3] is a drawing showing the remaining half of a circuitry drawing where the information reproducing device of the 1st example is detailed.

[Drawing 4] is a block diagram for explaining the function of the information reproducing device of the 1st example.

[Drawing 5] (A) and (B) are flow charts that show the operation in the voice registration mode and voice search mode of an information reproducing device of the 1st example, respectively.

[Drawing 6] (A) is a drawing showing the composition of the feature extracting part in drawing 4 and (B) is a drawing for explaining pattern matching processing.

[Drawing 7] (A) is a drawing for explaining search by pattern matching and (B) is a block diagram of the conventional sound reproduction device.

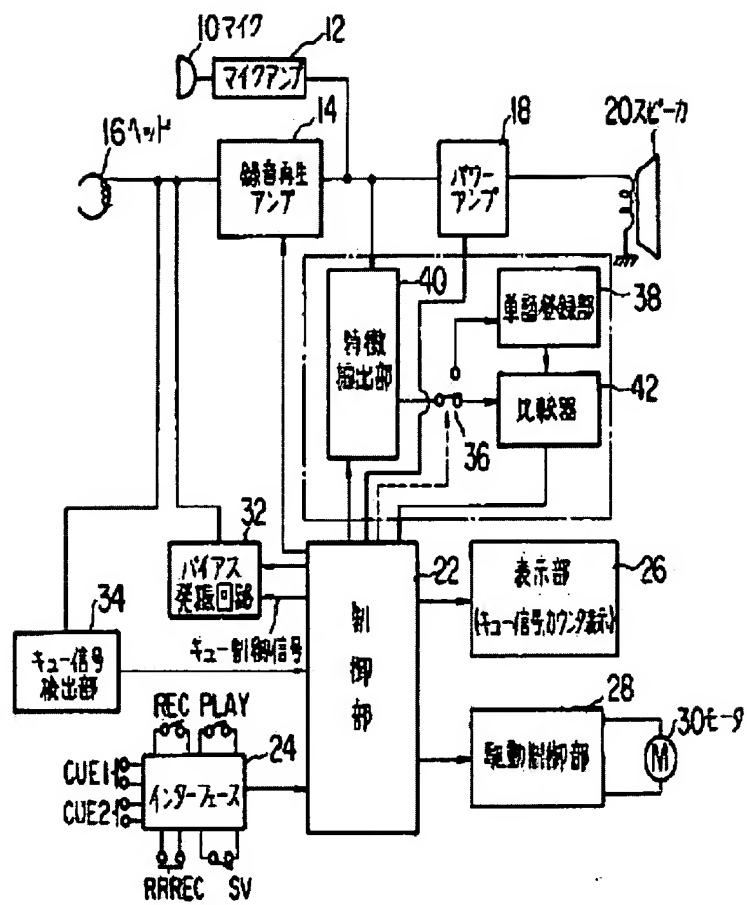
[Drawing 8] is a detailed circuitry drawing of the information reproducing device of the 2nd example.

[Drawing 9] is a block diagram for explaining the function of the information reproducing device of the 2nd example.

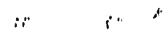
### **[Description of numbers]**

10... A microphone, 12... A microphone amplifier, 14... A sound recording playback amplifier, 16... A head 18... A power amplifier, 20... A loudspeaker, 22... A main control part, 24... An interface, 26... A display part, 28... A driving control part, 30... A motor, 32... A bias oscillating circuit, 34... A cue signal detecting part, 36... A switch, 38, 122... A word registering part (characteristic standard pattern registration means), 40... A feature extracting part (characteristic pattern extraction means), 42... A comparator (judging means)

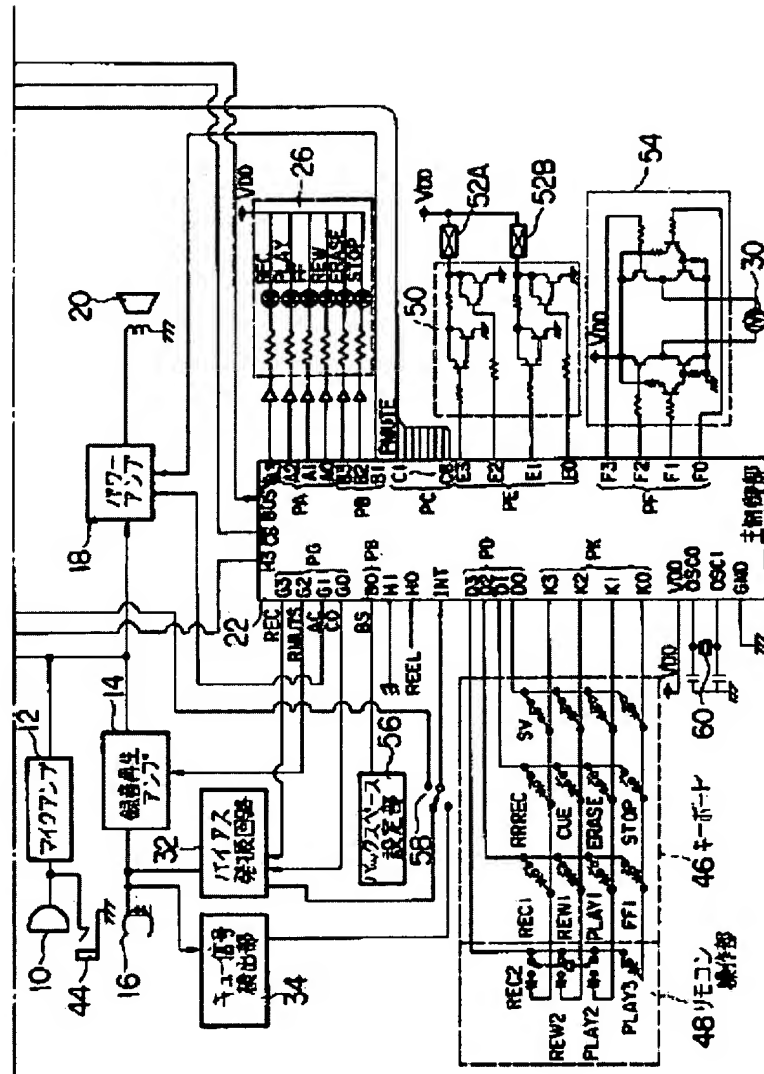
### Drawing 1



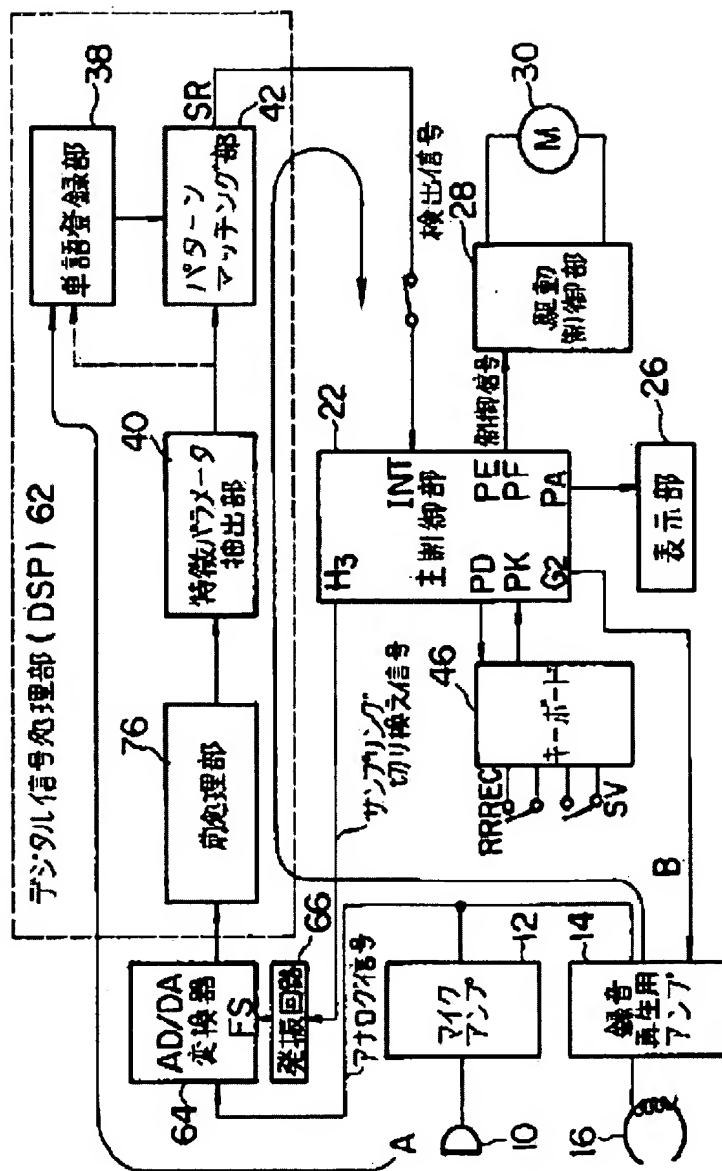
11. 12. 13.



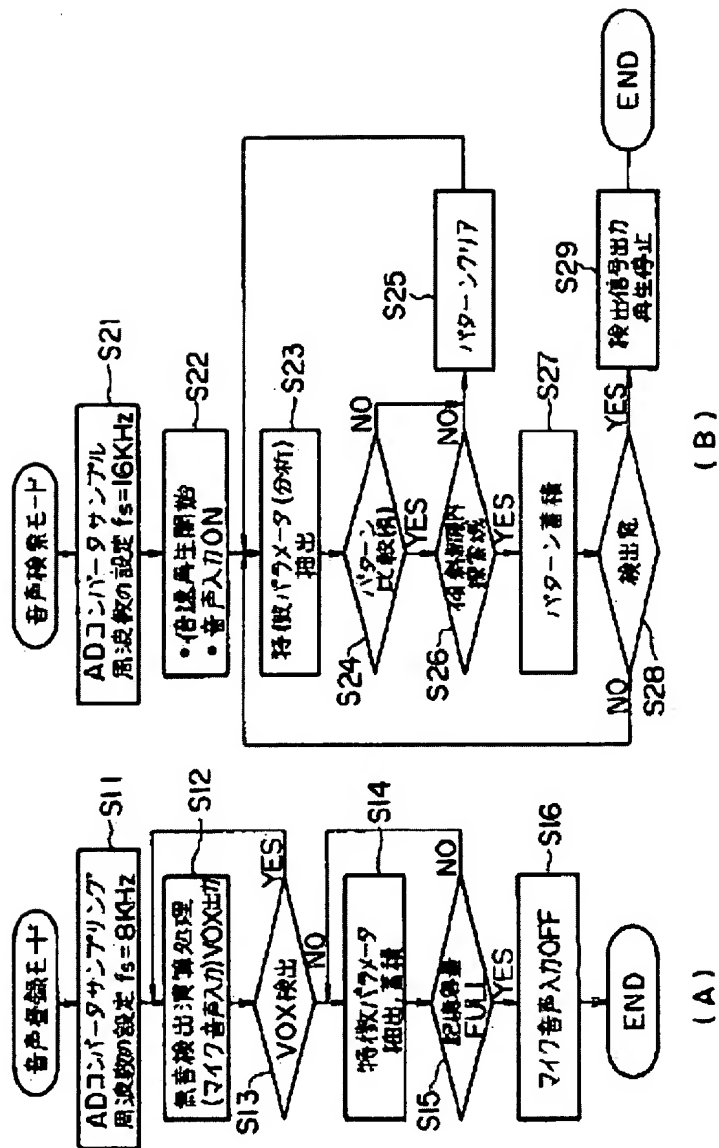
### Drawing 3



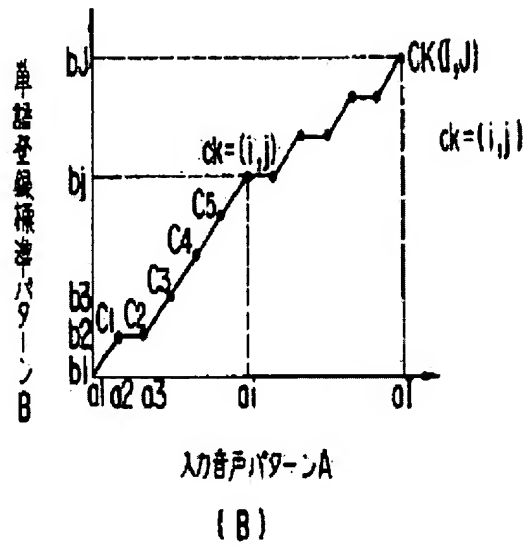
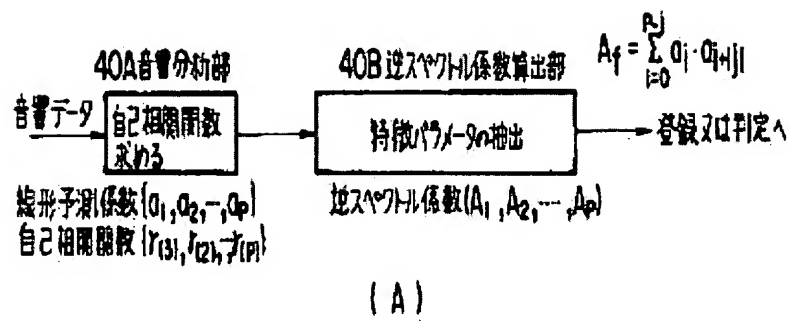
### Drawing 4



Drawing 5

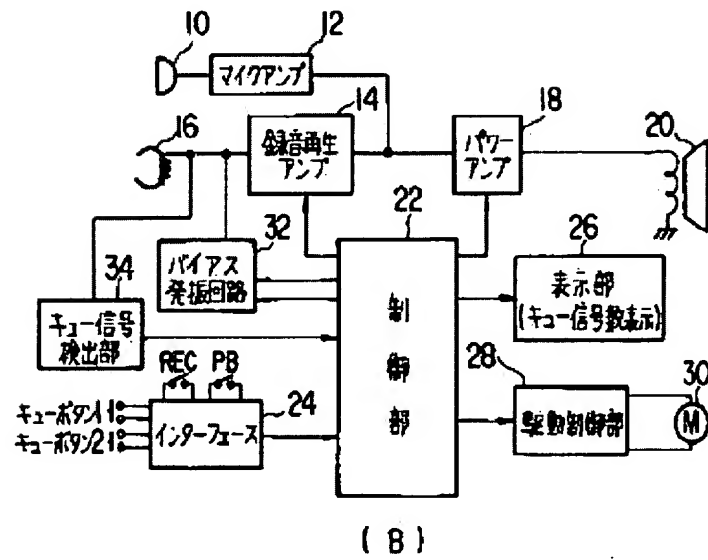
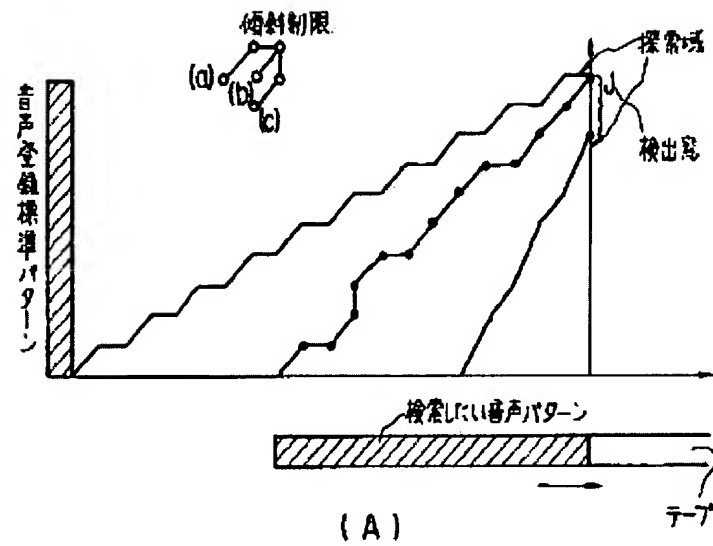


Drawing 6

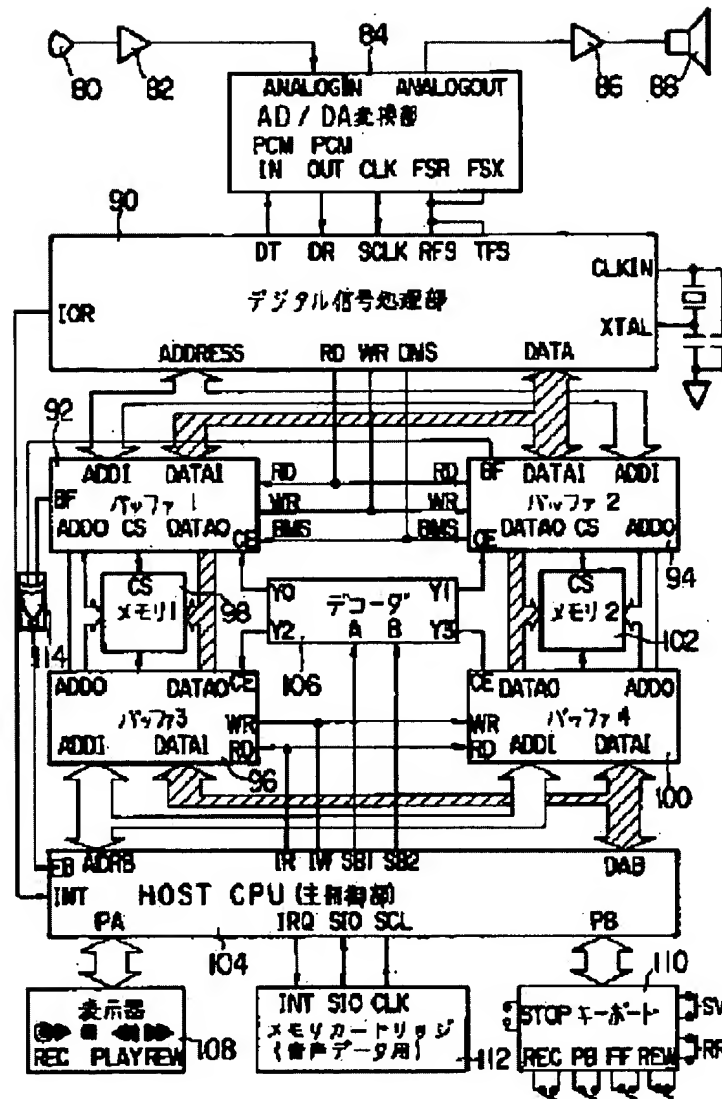




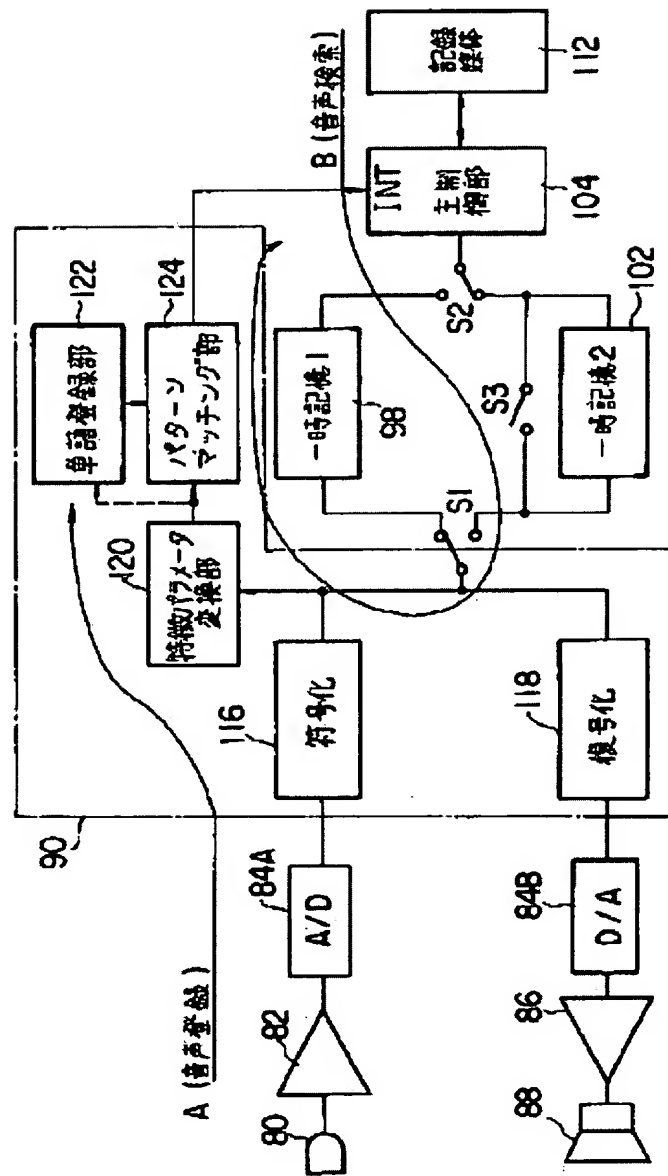
Drawing 7



Drawing 8



Drawing 9



### ABSTRACT OF REFERENCE 3

(11)Publication number :

06-139289

(43)Date of publication of application : 20.05.1994

(51)Int.Cl.

G06F 15/40

G10L 3/00

G11B 27/28

(21)Application number : 04-287705

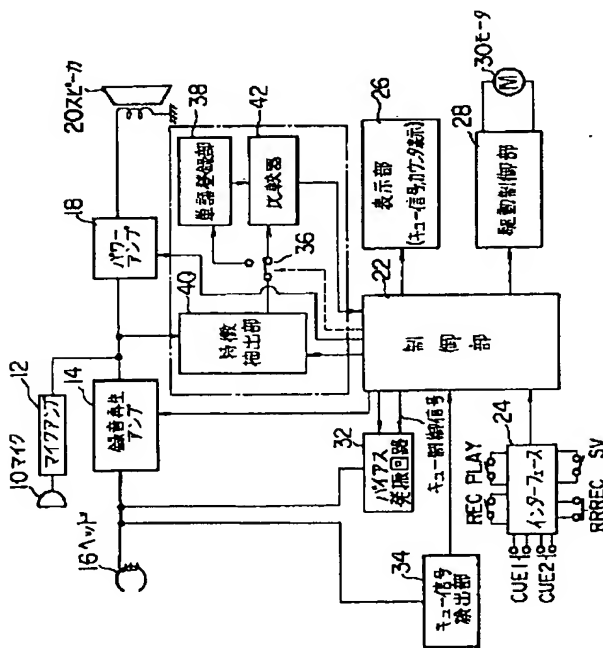
(71)Applicant : OLYMPUS OPTICAL CO LTD

(22)Date of filing :

26.10.1992

(72)Inventor : OKANO HIDEO

(54) INFORMATION REPRODUCING DEVICE



(57)Abstract:

**PURPOSE:** To retrieve a desired recording part in recorded voice information without operating any troublesome operation.

CONSTITUTION: In a voice registration mode, a feature extracting part 40 extracts a voice parameter from a voice signal

inputted from a microphone 10, and registers it through a switch 36 which is switched and controlled by a main control part 22 in a word registering part 38. In a retrieval mode, the main control part 22 controls a driving control part 28, and operates a high

speed reproducing operation. Then, the feature extracting part 40 extracts the voice parameter from a reproduction signal from a head 16, and transmits it through the switch 36 to a comparator 42. The comparator 42 compares the voice parameter registered in the word registering part 38 with the transmitted voice pattern, and transmits a detection signal to the main control part 22 at the time of detecting an equal part. The main control part 22 inputs the detection signal, then transmits a stop signal to the driving control part 28, and stops the high speed reproducing operation.

(11)特許出願公開番号

特開平6-139289

(43)公開日 平成6年(1994)5月20日

(51)Int.Cl. <sup>5</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/40	5 3 0 V	7218-5L		
G 1 0 L 3/00	5 5 1 G	9379-5H		
G 1 1 B 27/28		F 8224-5D		

審査請求 未請求 請求項の数 3 (全 18 頁)

(21)出願番号 特願平4-287705

(22)出願日 平成4年(1992)10月26日

(71)出願人 000000376

オリンパス光学工業株式会社  
東京都渋谷区幡ヶ谷2丁目43番2号

(72)発明者 岡野 秀生

東京都渋谷区幡ヶ谷2丁目43番2号 オリ  
ンパス光学工業株式会社内

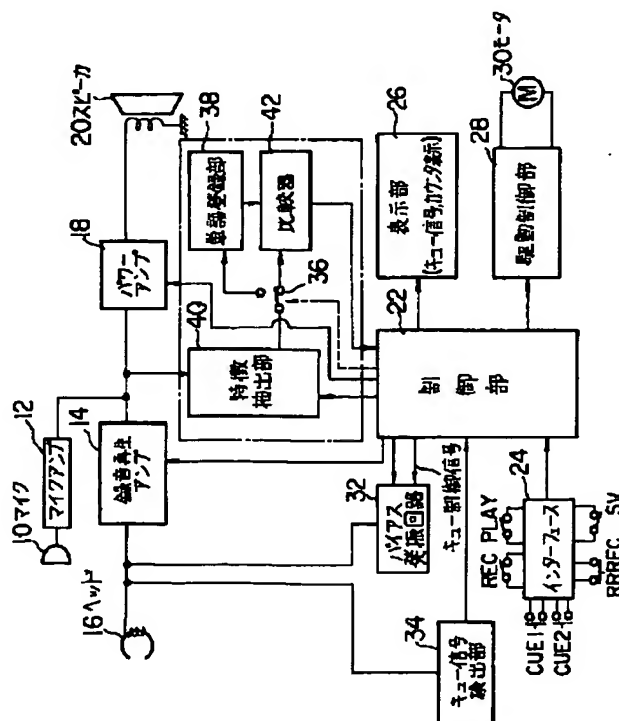
(74)代理人 弁理士 鈴江 武彦

(54)【発明の名称】 情報再生装置

(57) 【要約】

【目的】煩雑な操作を行わずに記録された音声情報中の  
所望の記録箇所を検索可能にする。

【構成】音声登録モードで、特徴抽出部４０は、マイク１０より入力された音声信号から音声パラメータを抽出し、これを主制御部２２により切り換え制御されるスイッチ３６を介して単語登録部３８に登録する。検索モードで、主制御部２２は駆動制御部２８を制御して高速再生動作を行う。そして、特徴抽出部４０は、ヘッド１６からの再生信号から音声パラメータを抽出し、スイッチ３６を介して、比較器４２に送る。比較器４２は、単語登録部３８に登録された音声パラメータと送られてきた音声パターンとを比較し、一致する箇所が検出できると、主制御部２２に検出信号を送る。主制御部２２は、この検出信号が入力されると、駆動制御部２８に停止信号を送り、高速再生動作を停止させる。



## 【特許請求の範囲】

【請求項1】 検索情報として利用する所定長さの音声情報の特徴的標準パターンを予め登録する特徴的標準パターン登録手段と、

上記所定長さの音声情報よりも長い音声情報が記録されている記録媒体より読み出された所定の音声信号を分析し特徴パターンを抽出する特徴パターン抽出手段と、抽出された特徴パターンと予め登録された所定長さの音声情報の特徴的標準パターンとを比較しその類似度を判定する判定手段と、  
10 を具備したことを特徴とする情報再生装置。

【請求項2】 デジタル化した音声信号を符号化し、この符号化されたデータを変換することで特徴パターンを抽出することを特徴とする請求項1に記載の情報再生装置。

【請求項3】 高速再生されたアナログ音声信号と所定長さの音声情報の特徴的標準パターンとを比較し、その類似度を判定できるように周波数軸を変換して高速再生信号を認識し検索することを特徴とする請求項1に記載の情報再生装置。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、音声を記録、再生する口述筆記装置や、電話回線や無線等で送られてきた音声  
を記録蓄積、再生する装置などに用いられる情報再生装置に関する。

## 【0002】

【従来の技術】最近、オフィスにおける事務合理化とともに特にマネージャ等の地位にあるものは多くの仕事量をこなさなければならない、これにほとんどの時間が費や  
30 されている。このため、手紙や報告書などの作成にあつては、その都度原稿を手書きすることなどできず、これに代わって原稿内容を予めテープに録音し、これを秘書が再生しながらタイプや筆記清書するようなことが行われている。

【0003】このような場合に用いられるものとして、例えば、特開昭60-25039号公報や特開昭63-112886号公報に開示されているような口述筆記装置がある。この口述筆記装置は、秘書がテープ内容を聞きながらタイプ操作を行うのに便利  
40 するように考えられており、このため多種多様にわたる諸機能が要求されている。

【0004】例えば、このような口述筆記装置において、口述者が筆記者に伝えようとするものに、予め記録されたメッセージの形式や優先度の有無、諸注意等を区別して伝えるようにしたいという要求がある。また口述筆記装置においてさらに要望される特徴として、記録媒体上に存在するメッセージの数、位置及び色々な形式が識別できるとともに、記録媒体上で所望のメッセージを容易に探索できるようにする  
50 というものがある。

【0005】そこで、従来、口述者は、例えば、テープレコーダ等の口述機を用いて後述した一連の文章の後に、他の後続する文章と区別するために、識別信号として、トーン又はキュー信号を記録するようにしていた。さらに、関連メッセージを記録する必要がある場合には、その記録の際に、上記識別信号とは異なるトーン又はキュー信号を記録するようにしていた。

【0006】識別信号を記録するための典型的な操作方は、口述者が口述による記録を行っているときに途中で  
10 筆記者に伝えたいメッセージ等が発生した場合、キュー釦を押し識別する信号としてキュー信号を記録し、また口述メッセージ（一連の文章）記録の終了を示すときに別のキュー釦を押し上記識別信号とは異なる識別信号を記録するというものである。

【0007】従って、筆記者は、筆記を行う前に、情報の高速再生を行って予め口述者の諸注意を聞くことができる。また、この高速再生時に、口述筆記装置の表示部には、記録媒体とキュー信号の相対的位置関係なども表示されるので、筆記者は聞きたい情報を視覚的に区別  
20 することができる。

【0008】このような口述筆記装置のシステム構成は、例えば図7の（B）に示すような構成をとる。なお、この図は、上記特開昭60-25039号の第1図を簡略化したものである。

【0009】即ち、録音動作においては、マイク10より入力された音声信号は、マイクアンプ12で増幅され、録音再生アンプ14を通してヘッド16によって不図示記録媒体に記録される。再生動作においては、記録媒体からヘッド16によって検出された信号は、録音再生アンプ14で増幅され、パワーアンプ18でさらに増幅されてスピーカ20に送られる。これらの一連の動作のコントロールは制御部22によって行われる。この制御部22は、また、インターフェース24を介して録音、再生のスイッチの状態を検出し、表示部26や駆動制御部28に制御信号を送る。駆動制御部28はモータ30の制御を行う。

【0010】録音時にキューボタンが押されると、インターフェース24から信号が制御部22に送られ、制御部22はこれに応じて、キュー制御信号をバイアス発振回路32に送り、ヘッド16を介してキュー信号を記録媒体に記録させる。再生または高速再生時には、キュー信号がヘッド16を介してキュー信号検出部34より検出され、これに応じて制御部22は表示部26又は駆動制御部28に制御信号を送り前述したような所定の動作を行わせる。

## 【0011】

【発明が解決しようとする課題】しかし、従来の口述筆記装置においては、口述記録された文章中の所望箇所を  
50 探す場合、口述者は、再生又は倍速再生を行い、実際に耳で聞いてその内容を確認しながら上記所望箇所を見つける

け出す操作を行わなければならない。一般に、このよう  
なときは、再生、早送り、巻き戻しの操作を繰り返し行  
わなければならない、所望の箇所を見つけ出すまでは煩  
雑な操作が必要になる。

【0012】本発明は、上記の点に鑑みてなされたもの  
で、記録された音声情報中の所望箇所を見つけ出すと  
き、再生、早送り、巻き戻しの操作を繰り返すといった  
煩雑な操作を行わずに、所望の箇所を検索することが可  
能な情報再生装置を提供することを目的とする。

【0013】

【課題を解決するための手段】上記の目的を達成するた  
めに、本発明による情報再生装置は、検索情報として利  
用する所定長さの音声情報の特徴的標準パターンを予め  
登録する特徴的標準パターン登録手段と、上記所定長さ  
の音声情報よりも長い音声情報が記録されている記録媒  
体より読み出された所定の音声信号を分析し特徴パター  
ンを抽出する特徴パターン抽出手段と、抽出された特徴  
パターンと予め登録された所定長さの音声情報の特徴的  
標準パターンとを比較しその類似度を判定する判定手段  
とを備えることを特徴とする。

【0014】

【作用】即ち、本発明の情報再生装置によれば、記録媒  
体に記録されている音声情報の所望の箇所を捜すため  
に、検索情報として利用する所定長さの音声情報の特徴  
的標準パターンを予め登録しておき、記録媒体より読み  
出された所定の音声信号を分析し特徴パターンを抽出し  
て、その抽出された特徴パターンと予め登録された所定  
長さの音声情報の特徴的標準パターンとを比較しその類  
似度を判定する。

【0015】

【実施例】以下、本発明の実施例を説明する。

(第1実施例) 図1は、本発明の第1実施例の構成を示  
す図で、同図において、従来と同様のものには、図7の  
(B)におけるのと同様の参照番号を付す。

【0016】即ち、録音動作においては、マイク10よ  
り入力された音声信号は、マイクアンプ12で増幅され  
た後、録音再生アンプ14を通してヘッド16によって  
不図示の記録媒体に記録される。また、再生動作にお  
いては、上記不図示記録媒体から上記ヘッド16によ  
って検出された信号が、上記録音再生アンプ14で増幅  
された後、パワーアンプ18でさらに増幅されてスピーカ  
20に送られる。これらの一連の動作は、主制御部22に  
よって制御される。この主制御部22はまた、インター  
フェース24を介して、録音スイッチREC、再生スイ  
ッチPLAY、キューボタンCUE1及びCUE2、音  
声登録ボタンRRREC、検索ボタンSV、等の状態を  
検出し、さらには表示部26や駆動制御部28に制御信  
号を送る。主制御部22からの上記制御信号に応じて、  
表示部26はキュー信号表示やカウンタ表示などの種々  
の表示を行い、また駆動制御部28はモータ30の駆動

制御を行う。

【0017】一方、上記録音時において、上記キューボ  
タンCUEが押されると、上記インターフェース24か  
ら信号が上記主制御部22に送られる。主制御部22は  
これに応じて、キュー制御信号をバイアス発振回路32  
に送り、上記ヘッド16を介してキュー信号を不図示記  
録媒体に記録させる。そして、再生又は高速再生時に  
は、キュー信号が上記ヘッド16を介してキュー信号検  
出部34によって検出され、その検出に応じて上記主制  
御部22は、表示部26又は駆動制御部28に制御信号  
を送り所定の動作を行わせる。以上は、従来の装置と同  
様である。本実施例の情報再生装置は、さらに次のよう  
な構成を有している。

【0018】即ち、音声登録ボタンRRRECが押され  
ると、上記インターフェース24から信号が上記主制御  
部22に送られる。主制御部22はこれに応じて、スイ  
ッチ36を単語登録部38側に切り換える。そして、上  
記マイク10より入力された音声信号は、マイクアンプ  
12を通して音声の特徴抽出部40に入力され、この特  
徴抽出部40で抽出された音声パラメータが、上記スイ  
ッチ36を介して上記単語登録部38に入力されて、そ  
こに登録される。

【0019】次に、検索ボタンSVが押されると、上記  
インターフェース24から信号が送られ、上記主制御部  
22は、高速再生動作信号を駆動制御部28に送り高速  
再生動作を行う。このとき、上記ヘッド16により不図  
示の磁気テープから再生信号が検出され、上記録音再生  
アンプ14で増幅した後、上記特徴抽出部40で音声パ  
ラメータが出力される。ここで、上記主制御部22が上  
記スイッチ36を切換制御することにより、比較器42  
にその音声パラメータが送られる。この比較器42は、  
上記単語登録部38に登録された音声パラメータと、こ  
の特徴抽出部40より送られてきた音声パターンとを比  
較し、一致する箇所が検出できると、上記主制御部22  
に検出信号を送る。主制御部22は、この検出信号が入  
力されると、上記駆動制御部28に停止信号を送り、高  
速再生動作を停止させる。

【0020】以上のようにして、倍速再生又は逆高速再  
生を行いながら、口述者が記録し直したい箇所を見つけ  
る際に、その希望の場所を、音声の登録によって見つ  
け出すことによって、再生、早送り、巻き戻しの操作を  
繰り返すといった煩雑な操作を行わずに、所望の音声記  
録箇所を検索することができ、キューマークの追加や録  
音した音声の編集を容易に行うことができるようになる。

【0021】ところで、上記不図示記録媒体として磁気  
テープを使用し、高速再生を行いながら検索を行う場  
合、高速再生で音声信号が変化してしまう。このため、  
高速再生した信号と登録した音声とが全く異なる音声  
信号になるため、高速再生された信号でも検索できる  
何らかの手法を盛り込む必要がある。そこで本実施例  
では、音



声登録のときのサンプリング周波数と高速再生のサンプリング周波数を変えることにより、登録した音声と高速再生された音声をパターンマッチングさせるときに時間軸を一致させるようにするものとする。具体的には、登録時点ではサンプリング周波数を8kHzで、倍速再生時は16kHzにする。このような手法を盛り込んだ本第1実施例の具体的な構成を図2及び図3に示す。

【0022】即ち、マイク10にマイクアンプ12を接続し、このマイクアンプ12に録音再生アンプ14を介して磁気ヘッド16を接続し、また録音再生アンプ14にパワーアンプ18を介してスピーカ20を接続している。なお、参照番号32はバイアス発振回路、44は外部マイク接続用端子である。

【0023】例えばマイクロコンピュータで構成される主制御部22は、録音再生アンプ14やパワーアンプ18に接続され、録音、再生の制御を行う。この主制御部22の出力ポートPDの各端子D0～D3と入力ポートPKの各端子K0～K3の間にキーボード46、フートスイッチやハンドコントローラなどのリモコン操作部48が接続されている。ここで、上記キーボード46は、録音キーREC1、巻き戻しキーREW1、再生キーPLAY1、早送りキーFF1、音声登録キーRRREC、キューキーCUE、消去キーERASE、停止キーSTOP、検索キーSVを有している。また、リモコン操作部48は、録音キーREC2、巻き戻しキーREW2、再生キーPLAY2、PLAY3、等を有している。

【0024】上記主制御部22の入出力ポートPAの各端子A0～A3及び入出力ポートPBの端子B2、B3には、表示回路26が接続されている。この表示回路26は、LEDにより録音、再生、巻き戻し、早送り、等の状態表示ができるように構成されている。

【0025】上記主制御部22の出力ポートPEの各端子E0～E3には、上記駆動制御部28の一部を成すソレノイド駆動回路50を介して、ソレノイド52A、52Bを接続している。これらソレノイド52A、52Bはメカ部を各モードに駆動するためのもので、ソレノイド52Aは録音再生モード時に駆動され、またソレノイド52Bは早送り巻き戻しモード時に駆動される。また、主制御部22の出力ポートPFの端子F0～F3には、上記駆動制御部28の一部を成すモータ駆動回路54を介して、モータ30を接続している。

【0026】上記主制御部22の入出力ポートPBの端子B0には、バックスペース設定部56を接続している。このバックスペース設定部56は、再生モードで一時停止し再び再生モードとした時に、所定の時間だけ巻き戻す時間を設定するためのものである。

【0027】上記主制御部22の出力ポートPGの各端子G0～G3の内、端子G0はバイアス発振回路32に接続しており、録音モード時に、上記キーボード46の

キューキーCUE操作に応じてCUE出力信号を発生する。また端子G1は、アラーム信号つまりAC出力を発生する。端子G2は、録音再生アンプ14に接続し、録音モードのとき、録音再生アンプ14の出力の立ち上がり所定の時間遅らせる信号を発生する。端子G3は、バイアス発振回路32に接続しており、上記キーボード46の録音スイッチREC1、REC2の操作に応じて、録音起動のREC出力信号を発生する。

【0028】上記主制御部22の入力ポートH0には、プリエンド検出のためのREEL入力を与えるようにしている。この場合のREEL入力は、不図示のリールの回転に応じたモータのパルス状回転信号が用いられている。また、入力ポートINTには、切換スイッチ58を接続し、このスイッチ58にキュー信号源としての上記バイアス発振回路32を接続するとともに、上記キュー信号検出部34を接続している。このキュー信号検出回路34は、再生モード時、磁気ヘッド16の再生出力よりキュー信号を検出するもので、この検出出力を上記入力ポートINTに入力するようにしている。そして、上記主制御部22の端子OSC0、OSC1には、発振回路60を接続している。

【0029】上記スイッチ58にはまた、デジタル信号処理部(DSP)62の端子SRが接続されている。一方、上記マイクアンプ12と録音再生アンプ14の結節点にAD/DA変換器64が接続されている。そして、このAD/DA変換器64と上記デジタル処理部62とは、それぞれ端子DRと端子POUT、端子DXと端子PIN、両CLK端子間、両FSX及びFSR端子間が接続され、デジタル化された音響データの送受信を行う。また、両CLK端子には、伝送速度を決めるクロックを発生する発振回路66が接続されている。

【0030】この発振回路66は、上記主制御部22の端子H3と接続されており、この端子H3から発生する信号で、クロック周波数を内部分周回路66Aにて変化させることができる。また、この発振回路66は、上記AD/DA変換器64のFS端子、及びタイミング制御回路68にも接続されている。このタイミング制御回路68は、上記デジタル処理部62及びAD/DA変換器64の端子FSX、FSRに接続され、伝送タイミングを制御している。

【0031】上記デジタル信号処理部62は、メモリ70にアドレス端子とデータ端子が接続されている。また、このデジタル信号処理部62の制御信号に接続されたデコーダ72は、選択信号を発生し、それを上記メモリ70の端子CSに入力するよう接続されている。

【0032】上記デジタル信号処理部62のデータ端子はまた、バッファ74に接続され、該バッファ74は上記デコーダ72とも接続されている。このバッファ74の端子D1～D8は、上記主制御部22の入出力ポートPCの端子C0～C7に接続されている。このバッファ

74は、処理スピードの異なる上記主制御部22と上記デジタル信号処理部62との間でデータの送受信を可能にするために用いられている。次に、このような構成における作用を、図4の処理の概略図を参照して説明する。

【0033】初めに、検索情報として利用する音声の特徴的標準パターンを予め登録する音声登録モードとして、図4中の経路Aを通る。また、この音声登録モードの処理の流れは、図5の(A)に示すフローチャートのようにになる。

【0034】まず、キーボード46の音声登録スイッチRRRECをONすると、主制御部22は、そのことをPDポートとPKポートより検知して、出力ポートH3から発振回路66のFT端子にサンプリング切り換え制御信号を入力する。この信号により、発振回路66は内部分周切換回路66Aにより、サンプリング周波数を8KHzに設定する(ステップS11)。

【0035】そして、マイク10から入力された音響信号は、マイクアンプ12で増幅された後、AD/DA変換器64に入力される。このAD/DA変換器64は、サンプリング周波数8KHzで入力アナログ音響信号をサンプリングしてデジタル信号に変換する。このデジタル信号化された音響信号は、デジタル信号処理部62の内部の前処理部76に送られる。この前処理部76では、入力した音響データを高域強調フィルタ(プリエンファシス)に通し、適応フィルタなどを利用して雑音除去を行った後、マイク10から入力された音響信号が音声であるかをスペクトラムやレベルなどより判定し、もし無音(VOX)と判定するとVOX信号を出力する(ステップS12)。このVOX信号が検出されなくなると(ステップS13)、音響データは10msec～30msec毎のブロックデータにまとめられて特徴パラメータ抽出部40に送られる。

【0036】この特徴パラメータ抽出部40内部の処理は、図6の(A)に示すようになる。即ち、前処理部76より送られてくるブロック音響データは、音響分析部\*

$$s'(n) = -a_1 s(n-1) - a_2 s(n-2) - \dots - a_p s(n-p)$$

【0039】この式は、 $s(n)$ はそれよりも過去の値の線形結合で表していることから、線形予測と呼ばれている。この線形予測係数は予測値 $s'(n)$ ができるだけ

$$\begin{aligned} \varepsilon(n) &= s(n) - s'(n) \\ &= s(n) + a_1 s(n-1) \\ &\quad + a_2 s(n-2) \dots + a_p s(n-p) \end{aligned}$$

【0040】 $\varepsilon^2(n)$ の時刻nについての総和 $y^2 = \varepsilon^2(1) + \varepsilon^2(2) + \varepsilon^2(3) + \dots$ が最小になるように係数を決めるわけである。その条件を求めると、

\*40Aに入力される。この音響分析部40Aは、線形予測分析から自己相関関数を計算する。このとき、線形予測係数 $\{a_1, a_2, a_3, \dots, a_p\}$ と自己相関関数 $\{r(1), r(2), r(3), \dots, r(p)\}$ が求まり、これを用いて逆スペクトル係数算出部40Bがブロックデータ毎に特徴パラメータとしての逆スペクトル係数 $\{A_1, A_2, A_3, \dots, A_p\}$ を求め蓄積する(ステップS14)。この逆スペクトル係数は、単語登録部38に、その記憶容量が一杯になるまで(数秒分)検索データとして蓄積される(ステップS15)。そして、記憶容量が一杯になったならば、マイク10からの音声入力信号の受け付けを終える、即ち無視する(ステップS16)。ここで、線形予測法(線形予測分析)について説明を加える。

【0037】線形予測法(linear prediction method)による分析のことを、通常、LPC分析といっている。LPCとは、線形予測係数(linear prediction coefficients)のことで、スペクトルに関する情報が10個程度の係数に集約されている。この方法の特徴は、音源で発声した声の元になる音が声道を通過して口から出てくるまでの音声の生成過程を音声波形から求めてしまうものであり、単にスペクトルの分析法としてだけでなく、音声合成法としても最も有効な手段となっている。

【0038】連続した音声波形 $s(t)$ を $\Delta t$ 秒毎(通常は $100\mu s \sim 125\mu s$ が多い)にN個(150個～300個程度)を読み取る音声波形が $\{s(\Delta t), s(2\Delta t), \dots, s(N\Delta t)\}$ という数値データとして得られる。このような波形を読み取る操作のことをサンプリングと称し、読み取り間隔 $\Delta t$ のことをサンプリング周期と称する。複雑化を避けるため $\Delta t$ を省略して、時刻 $t = n\Delta t$ における波形データを $s(n)$ と書く。ある時刻n(実際には時刻は $n\Delta t$ )での値 $s(n)$ をそれより過去の値をp個だけ使って次のような予測すると考える(ここで、 $s'(n)$ を $s(n)$ に対する予測値として考える)。

\* $s(n)$ に近い値になるように決める。その有効方法として最小2乗法が一般的に用いられている。即ち、いま予測誤差を $\varepsilon(n)$ とすれば、

【0041】

【数1】

$$\begin{array}{rcl}
 & 9 & 10 \\
 r(0)a_1 + r(1)a_2 + r(2)a_3 + \cdots + r(p-1)a_p & = & -r(1) \\
 r(1)a_1 + r(0)a_2 + r(1)a_3 + \cdots + r(p-2)a_p & = & -r(2) \\
 \vdots & & \vdots \\
 r(p-1)a_1 + r(p-2)a_2 + r(p-3)a_3 + \cdots + r(0)a_p & = & -r(p)
 \end{array}
 \left. \vphantom{\begin{array}{rcl} & 9 & 10 \\ r(0)a_1 + r(1)a_2 + r(2)a_3 + \cdots + r(p-1)a_p & = & -r(1) \\ r(1)a_1 + r(0)a_2 + r(1)a_3 + \cdots + r(p-2)a_p & = & -r(2) \\ \vdots & & \vdots \\ r(p-1)a_1 + r(p-2)a_2 + r(p-3)a_3 + \cdots + r(0)a_p & = & -r(p) \end{array}} \right\}$$

これは正規方程式と呼ばれて、 $r(0)$ 、 $r(1)$ 、… \* 【0042】  
は信号  $s(n)$  の自己相関関数と呼ばれる。この自己相  
関関数は、 \* 【数2】

$$r(m) = \sum_{n=1}^{N-m} s(n)s(n+m) \quad m=0, 1, 2, \dots, p$$

【0043】が定義式で、これは  $\{s(n)\}$  自身とそれ  
を  $m$  サンプルだけ全体をずらしたもののどうしを掛けて  
加算したものである。 $m=0$  のときには、 $r(0)$  は信  
号の2重和であるから  $s(n)$  のエネルギーである。従  
て、サンプル値  $s(1)$ 、 $s(2)$ 、…、 $s$   
( $N$ ) から上式の自己相関関数を計算して、それで作※

$$\sigma^2 = r(0)A_0 + 2r(1)A_1 + 2r(2)A_2 + \cdots + 2r(p)A_p$$

【0045】次に、登録された音声を検索する音声検索  
モードにする。この音声検索モードは、図4中の経路B  
をとる。また、この音声検索モードの処理の流れは、図  
5の(B)に示すフローチャートのようにになる。

【0046】まず、キーボード46の検索キーSVをO  
Nすると、主制御部22はそれをPDポートとPKポ  
ートより検知して、出力ポートH3より発振回路66のF  
T端子にサンプリング切り換え制御信号を入力する。これ  
により、高速再生に対応したサンプリング周波数に設  
定される。この実施例では倍速再生で検索とすると、1  
6KHzに設定される(3倍速では24KHzとなる)  
(ステップS21)。また、主制御部22は、PEポ  
ートとPFポートに高速再生動作信号を出力し、例えば倍  
速で再生を行うようにする(ステップS22)。

【0047】そして、ヘッド16から入力された音響信  
号は録音再生アンプ14で増幅された後、AD/DA変  
換器64に入力される。AD/DA変換器64は、サン  
プリング周波数16KHzでサンプリングを行ってアナ  
ログ音響信号をデジタル信号に変換する。このデジタル  
信号化された音響信号は、デジタル信号処理部62の内  
部の前処理部76に送られる。この前処理部76で、音  
響データ区間は10msec~30msec毎のブロッ  
クデータにまとめられ、高域強調フィルタ(プリエンフ  
ァンス)に通され、適応フィルタなどを利用して雑音除  
去を行った後、特徴パラメータ抽出部40に送られる。  
この特徴パラメータ抽出部40では、音声登録モードの  
ときと同様に、この前処理部76から送られてくるプロ

※られる連立方程式を解いて線形予測係数を求める。

【0044】また、逆スペクトル係数は、この自己相関  
関数  $\{r(0), r(1), \dots, r(p)\}$  と線形予測  
係数  $\{a_1, a_2, \dots, a_p\}$  と予測誤差の総和の  $\sigma^2$  か  
ら次の式から定まる。

ック音響データ信号  $Z(n)$  が音響分析部40Aに入力  
されて、線形予測分析から自己相関関数が計算される。  
このとき、線形予測係数  $\{a_1, a_2, a_3, \dots,$   
 $a_p\}$  と自己相関関数  $\{r(1), r(2), r$   
(3), ...,  $r(p)\}$  が求まり、次に、逆スペクトル  
係数算出部40Bでこれを用いて逆スペクトル係数  $\{A$   
 $_1, A_2, A_3, \dots, A_p\}$  を求め(ステップS2  
3)、今度は、それを比較器としてのパターンマッチン  
グ部42に送る。

【0048】パターンマッチング部42においては、標  
準音声のパターンと検索する音声パターンとを照合する  
ため、パターン同士のスペクトラムの形を評価するため  
スペクトル距離を評価する。この距離尺度としては、L  
PCケプストラム距離尺度や重み付け尤度比(WLR)  
尺度最尤スペクトル距離尺度法などがある。この実施例  
では、単語登録部38の標準音声パターンと再生音声パ  
ターンのそれぞれ予測残差(予測誤差)の2乗和  $\sigma^2$  を  
求め、この予測残差より尤度  $l = -\log \sigma^2$  で導いた  
値を類似度として使用し、再生音声パターンと標準音声  
パターンのスペクトル距離を求めパターンマッチングを  
行う。図6の(B)はこのパターンマッチングの処理を  
示す図である。ここで、入力パターンAは10msec  
~30msecに区間分けされたブロックデータの特徴  
ベクトルとして  $\{a_1, a_2, a_3, \dots, a_I\}$  で表さ  
れ、単語登録部38の標準音声パターンBは  $\{b_1, b$   
 $_2, b_3, \dots, b_J\}$  で表わす。これらパターンA、B  
からなる平面を考えたとき、この平面上の格子点C=

(i, j) の系列は

$$F = C_1, C_2, C_3, \dots, C_k, \dots, C_k \quad C_k = (i_k, j_k)$$

【0049】となる。また、a i と b j のスペクトル距離を  $d(i, j) = d(i, j)$  とする。パターンマッチングでは、このスペクトル距離の最小のものを探索するが、ここでは簡略化のために、パターンマッチングは動\*

\* 的計画法のように考え累積距離 g を求める問題に変える。

【0050】

【数3】

$$g(C_k) = g(i, j)$$

$$= \begin{bmatrix} g(i-2, j-1) + 2d(i-1, j) + d(i, j) \\ g(i-1, j-1) + 2d(i, j) \\ g(i-1, j-2) + 2d(i, j-1) + d(i, j) \end{bmatrix}$$

で計算が繰り返行われる。

【0051】このように標準パターンと再生音声を一フレーム(10msec~30msecのブロックデータ単位)でシフトしながらパターンマッチングし、距離が一定の条件を満たすとき、対応する単語が存在したことを判定する。

【0052】即ち、入力パターンに一致する標準パターンがなければ(ステップS24)、入力パターンをクリアした後(ステップS25)、上記ステップS23へ戻り、また一致するパターンがあれば、図7の(A)に示すような検索領域と傾斜制限を設け、マッチングの判定を行う(ステップS26)。そして、この検索領域と傾斜制限から外れていれば、上記ステップS25へ進み、外れていなければ、そのパターンを蓄積、即ち累積距離 g を求める(ステップS27)。そして、累積距離 g が検索窓に到達していなければ(ステップS28)、上記ステップS23に戻り、検索窓に達したと判断されたときには、パターンマッチング部42より主制御部22のINT端子に検出信号が送られ、これにより主制御部22は再生停止を行うためPEポートとPFポートより駆動制御部28に制御信号を送る(ステップS29)。また検索中には、主制御部22のPAポートより表示部26に再生表示を点滅させるような表示を成させるための制御信号が送られる。

【0053】以上のように、音声メッセージ及び音声メッセージ識別信号を後から記録、追加する場合に口述者の記録したい箇所を見つけ出すとき、希望の場所に記録されている音声に相当する音声の登録を行って、倍速再生または逆高速再生を行いながら希望の場所を音声認識を行うことで、音声の記録によるテープ位置で見つけ検索することができる。また、今までキュー信号をわざわざ入れて秘書へ伝言や指示を行っていたものが、例えば「秘書宛」と録音中にマイクロホンから音声を入力することによって、その入力した音声を頼りに所望箇所の検索することが可能になる。

【0054】なお、以上の説明では、高速再生された信号でも検索できるように、音声登録のときと高速再生の

ときのサンプリング周波数を変える手法を用いるものとして説明したが、例えば次のような手法を用いることも可能である。

(1) パラメータ抽出前に単語登録音声または再生信号の音程変換処理(周波数変換処理)を行う。

(2) パターンマッチング時の周波数軸変換演算処理を行う。

(3) 高速再生時にはゲインが上がるため利得補正を行う。

(第2実施例)次に、本発明をデジタル録音再生装置に適用した場合を説明する。

【0055】即ち、本第2実施例は、音声デジタル信号変換した後、デジタル信号処理によって音声を符号化し音声の特徴パラメータを抽出してテープやディスクや半導体メモリ等に記憶する。そして、音声検索は前記符号化のときに求まるパラメータを変換し、音声のパターンマッチングを行い音声の検索を行う。

【0056】図8において、まずマイク80から入力された音声信号は、第1アンプ82に入力される。この第1アンプ82の出力端子は、AD/DA変換部84の端子ANALOG INに接続されている。また、AD/DA変換部84の端子ANALOG OUTと第2アンプ86の入力端子が接続されていて、そして、この第2アンプ86の出力端子にスピーカ88が接続されている。つまり、デジタル化されている音声信号は、AD/DA変換部84にてアナログ信号に変換して第2アンプ86に送られ、該アンプ86で増幅した後、スピーカ88に出力される。

【0057】AD/DA変換部84とデジタル信号処理部90間は、端子PCMINと端子DT、端子PCMOUTと端子DR、端子CLKと端子SCCLK、端子FSRや端子FSXと端子RFSや端子TFSとがそれぞれ接続されていて、AD/DA変換部84でデジタル変換された信号をデジタル信号処理部90に転送したり、デジタル信号処理部90で処理された信号をAD/DA変換部84に転送を行う。

【0058】上記デジタル信号処理部90の端子ADD

RESSは、第1バッファ92の端子ADDIと第2バッファ94の端子ADDIに接続されている。また、デジタル信号処理部90の端子DATAは、第1バッファ92の端子DATAIと第2バッファ94の端子DATAIに接続されている。デジタル信号処理部90のそれぞれ端子RD, WR, DMSは、第1バッファ92の端子RD, WR, BMSと第2バッファ94の端子RD, WR, BMSに接続されている。

【0059】第1バッファ92の端子ADD0と第3バッファ96の端子ADD0と第1メモリ98のADDR ESS端子とが接続され、第1バッファ92の端子DATA0と第3バッファ96の端子DATA0と第1メモリ98のDATA端子とが接続されている。同様に、第2バッファ94の端子ADD0と第4バッファ100の端子ADD0と第2メモリ102のADDRESS端子とが接続され、第2バッファ94の端子DATA0と第4バッファ100の端子DATA0と第2メモリ102のDATA端子とが接続されている。また、上記第1バッファ92と第3バッファ96の各CS端子と第1メモリ98のCS端子とが接続され、第2バッファ94と第2メモリ102のCS端子とが接続されている。

【0060】主制御部104の端子ADRBは、第3バッファ96の端子ADDIと第4バッファ100の端子ADDIに接続されている。また、主制御部104の端子DABは、第3バッファ96の端子DATAIと第4バッファ100の端子DATAIに接続されている。主制御部104のそれぞれ端子IR, IWは、第3バッファ96の端子RD, WRと第4バッファ100の端子RD, WRに接続されている。

【0061】主制御部104の端子SB1と端子SB2は、デコーダ106の端子Aと端子Bに接続されている。このデコーダ106の端子Y0は、第1バッファ92の端子CEに接続されている。同様に、デコーダ106の端子Y1は第2バッファ94の端子CEに、端子Y2は第3バッファ96の端子CEに、端子Y3は第4バッファ100の端子CEにそれぞれ接続されている。

【0062】また、上記デジタル信号処理部90のIOR端子は、主制御部104のINT端子に接続されている。主制御部104のPAポート端子は、表示器108に接続されていて、各動作状態を表示する。また、主制御部104のPBポート端子は、キーボード110と接続されている。そして、主制御部104とメモリカートリッジ112の間は、端子IRQと端子INT、端子SIOと端子SIO、端子SCLと端子CLKとが接続されている。

【0063】また、第1バッファ92と第2バッファ94の端子BFは、論理回路114に接続されている。そして、この論理回路114出力端子と主制御部104の端子FBが接続されている。次に、このような構成のデ

ジタル録音再生装置の動作を説明する。

【0064】初めに、録音モードでは、キーボード110のRECスイッチが押されると、主制御部104は、PBポートを通じて録音開始を検知して、第3バッファ96と第1バッファ92を通じてデジタル信号処理部90に圧縮処理準備をさせる命令コードを転送する。

【0065】ここで、マイク60より入力した音声信号は、第1アンプ61で増幅された後、AD/DA変換部84に送られ、該AD/DA変換部84でアナログ/デジタル変換される。このAD/DA変換部84にてデジタル変換された信号は、デジタル信号処理部90に転送される。デジタル信号処理部90では、デジタル化した音声信号を10msec~30msecのブロックデータにまとめた後、このブロックデータからCELP符号化方式のような分析形音声符号化方式を利用して音声の特徴パラメータを抽出し、パラメータの量子化を行う。

【0066】一方、主制御部104は、端子SB1とSB2より制御信号をデコーダ106に送ることにより予め第1バッファ92のCE端子に信号を送り、この第1バッファ92がデータを受け付けられるようにする。

【0067】そして、デジタル信号処理部90は、データ端子を通じて符号化したデータを第1バッファ92に転送する。第1バッファ92は、第1メモリ98にこのデータを転送した後、端子BFよりメモリにデータを転送したことを伝える信号を主制御部104のEB端子に送る。

【0068】これに応じて、主制御部104は、端子SB1, SB2の信号を送ることにより、デコーダ106の端子Y1, Y3より第2バッファ94と第4バッファ100のCE端子に制御信号を送り、データ入出力が可能な状態にする。

【0069】そして、第3バッファ96を通じて第1メモリ98の符号化データを主制御部104が受取り、これをシリアル端子IRQ, SIO, SCLを通じてメモリカートリッジ112に転送する。

【0070】主制御部104が上記の作業を行っている間に、デジタル信号処理部90は、次の入力信号の音声信号ブロックの演算を行い、符号化したデータを第2バッファ94に転送する。

【0071】第2バッファ94から第2メモリ102に符号化データが転送されると、BF端子から論理回路114に制御信号を出力し、該論理回路114の出力端子から主制御部104のEB端子にデータ転送されたことを伝える。

【0072】主制御部104は、これに応じて、端子SB1, SB2の信号を送ることにより、デコーダ106の端子Y0, Y2より第1バッファ92と第3バッファ96のCE端子に制御信号を送り、データ入出力が可能な状態にする。

【0073】その後は、キーボード110のSTOPボ

タンが押され、主制御部104がそれを認識してバッファを通じてデジタル信号処理部90に停止信号を送るまで、上記の操作を同様に繰り返し行う。即ち、デジタル信号処理部90と主制御部104との間のデータ転送は、第1バッファ92、第1メモリ98、第3バッファ96による経路と、第2バッファ94、第2メモリ102、第4バッファ100による経路との2つの経路を交互に介して行われる。

【0074】このように、符号化されたデータは、CELPのような分析符号化を利用したものは線形予測係数がデータとして蓄積されているため、パターンマッチングのための音声の特徴パラメータに変換することが可能で、音声検索も前述した第1実施例のテープレコーダと比較して高速に検索できるところに特徴をもつ。

【0075】図9には、本実施例のデジタル録音再生装置の機能ブロック図である。同図中の経路Aに示す流れは音声登録モードであり、経路Bは音声検索モードの流れを示している。

【0076】音声録音モードでは、まずマイク80より音声を入力して上記第1アンプ82に相当するマイクアンプで増幅された音声信号はAD/DA変換部84内に構成されたA/D変換器84Aでデジタル信号に変換された後、デジタル信号処理部90に送られる。このデジタル信号処理部90は、同図に示すような構成をとり、まず、符号化部116に信号が送られ、前記の分析符号化される。そして、この符号化されたデータは、上記バッファ及びデコーダで構成されるスイッチS1～S3により上記メモリ98、102に相当する一時記憶部を介して主制御部104に送られ、上記メモリカートリッジ112に相当する記録媒体に記録される。そして、通常、この記録媒体には、単語登録部122に登録される音声情報よりもはるかに長い即ち多くの音声情報が記録される。

【0077】音声再生モードでは、記録媒体112からの符号化データが主制御部104、スイッチS1～S3、一時記憶部98、102を介して、デジタル信号処理部90の復号化部118に送られて、復号される。そして、復号された音声信号AD/DA変換部84内に構成されたD/A変換器84Bでアナログ信号に変換された後、上記第2アンプ86に相当するパワーアンプで増幅された後、スピーカ88に出力される。

【0078】音声登録モードでは、マイク80より音声を入力して上記第1アンプ82に相当するマイクアンプで増幅された音声信号はAD/DA変換部84内に構成されたA/D変換器84Aでデジタル信号に変換された後、デジタル信号処理部90に送られる。このデジタル信号処理部90では、まず、符号化部116に信号が送られ、前記の分析符号化される。この符号化したデータの特徴パラメータ変換部120に送り、データの中に含まれる線形予測係数を利用して特徴パラメータに変換し

て上記第1実施例と同様な逆スペクトル係数を得て、単語登録部122に数秒間登録される。この時、上記第1実施例と同様に音声起動によって登録が行われる。

【0079】また、音声検索モードでは、主制御部104より記録場所が管理された記録媒体（メモリカートリッジ）112からの符号化データが、スイッチS1～S3と一時記憶部122A、112Bを介して、特徴パラメータ変換部120に転送され、前記と同様に特徴パラメータに変換され、パターンマッチング部124に送られる。このパターンマッチングの方法は上記第1実施例に示すような例と同様な方法を採用。パターンマッチングが検出されると、主制御部104のINT端子に検出信号が送られ、記録場所を見つけたとして、所定の停止処理を行う。

【0080】デジタル録音再生装置では、符号化されたデータに線形予測のパラメータが含まれているためこれを利用し、線形予測係数からパターンマッチングのための音声の特徴パラメータに変換することが可能であり、音声検索も上記第1実施例のテープレコーダと比較して演算量が少ないことや、ランダムアクセスによる高速に検索できる効果がある。

【0081】なお、本発明の考え方を推し進めれば、音声認識された特定の言葉を消去したり、別の言葉に置換するといった応用も可能である。例えば、消去の場合には、消去したい言葉を登録し、その言葉が含まれている部分の再生中に消去釦を押すことにより、自動的に、登録している言葉に対応する言葉の消去ができる。従って、消去したい言葉が多い場合には、非常に有効である。

【0082】

【発明の効果】以上詳述したように、本発明の情報再生装置によれば、記録された音声情報の所望の箇所を見つけ出すとき、希望の場所を音声の登録によって見つけ出すことができるので、再生、早送り、巻戻しの操作を繰り返すといった煩雑な操作を行わずに上記所望の箇所を検索することができる。

【図面の簡単な説明】

【図1】第1実施例の情報再生装置のブロック構成図である。

【図2】第1実施例の情報再生装置の詳細な回路構成図の半分を示す図である。

【図3】第1実施例の情報再生装置の詳細な回路構成図の残りの半分を示す図である。

【図4】第1実施例の情報再生装置の機能を説明するためのブロック構成図である。

【図5】(A)及び(B)はそれぞれ第1実施例の情報再生装置の音声登録モード及び音声検索モードにおける動作を示すフローチャートである。

【図6】(A)は図4中の特徴パラメータ抽出部の構成を示す図であり、(B)はパターンマッチング処理を説

明するための図である。

【図7】(A)はパターンマッチングによる検索を説明するための図であり、(B)は従来の音声再生装置のブロック構成図である。

【図8】第2実施例の情報再生装置の詳細な回路構成図である。

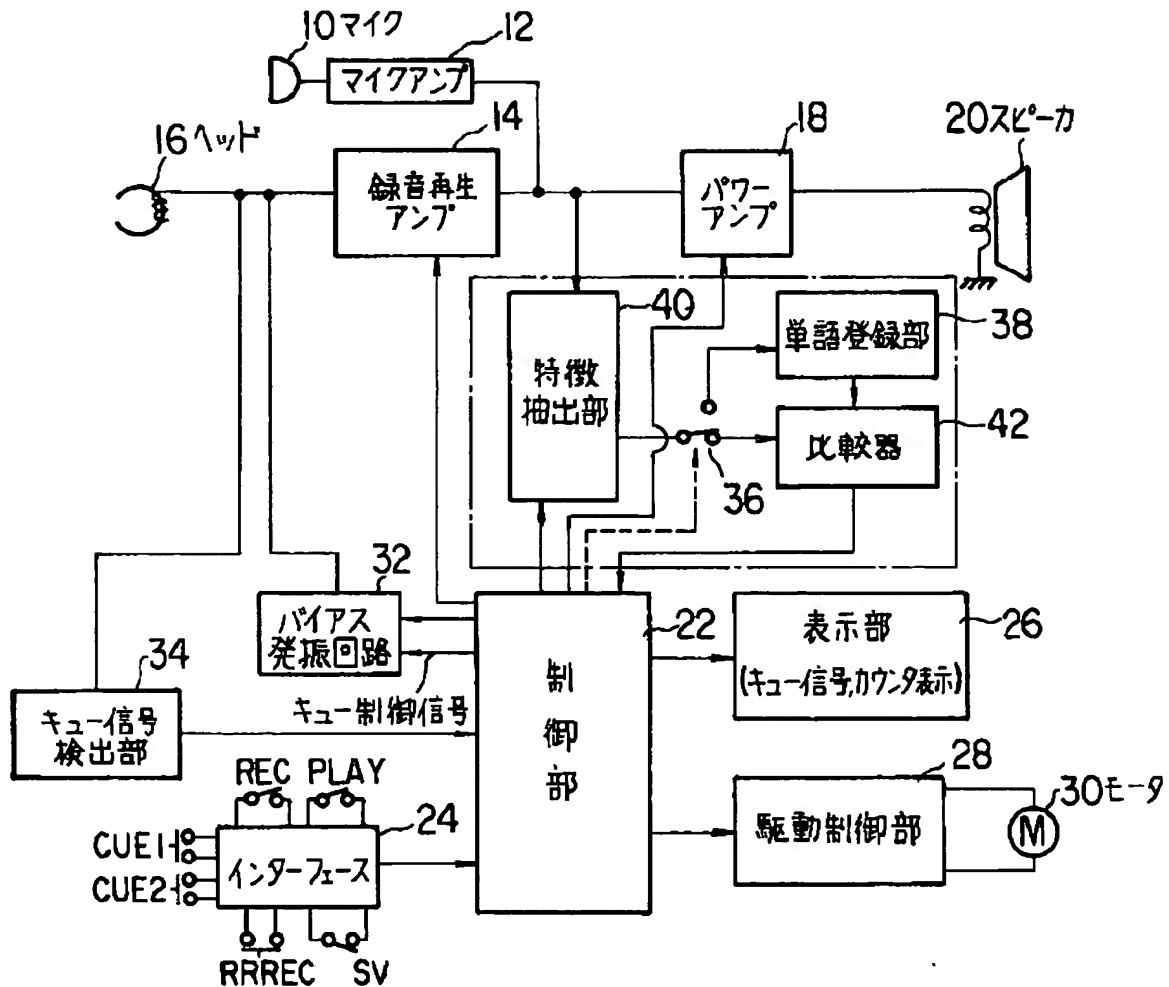
【図9】第2実施例の情報再生装置の機能を説明するためのブロック構成図である。

【符号の説明】

\*

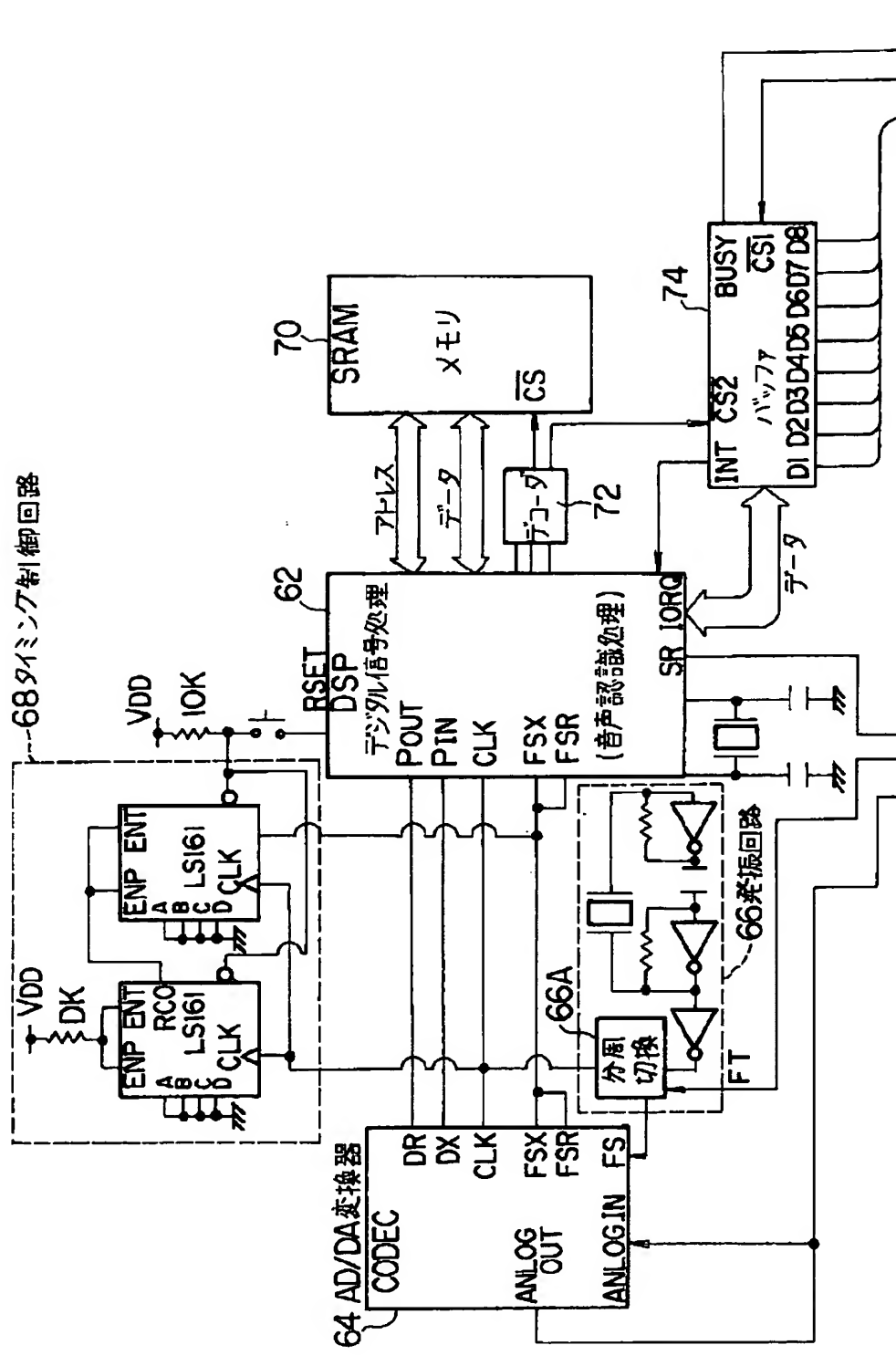
\* 10…マイク、12…マイクアンプ、14…録音再生アンプ、16…ヘッド、18…パワーアンプ、20…スピーカ、22…主制御部、24…インターフェース、26…表示部、28…駆動制御部、30…モータ、32…バイアス発振回路、34…キュー信号検出部、36…スイッチ、38、122…単語登録部（特徴的標準パターン登録手段）、40…特徴抽出部（特徴パターン抽出手段）、42…比較器（判定手段）。

【図1】



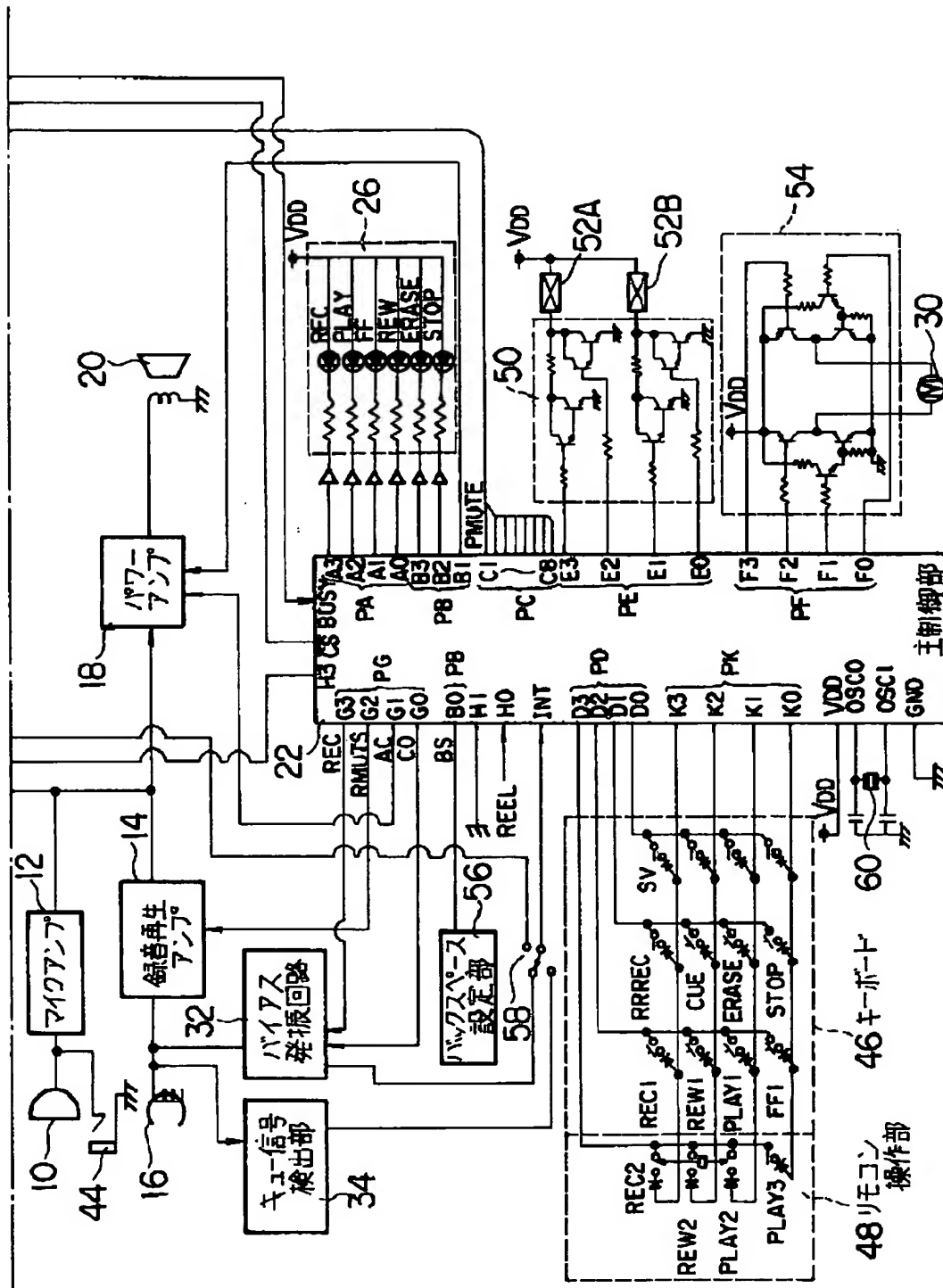


【図2】

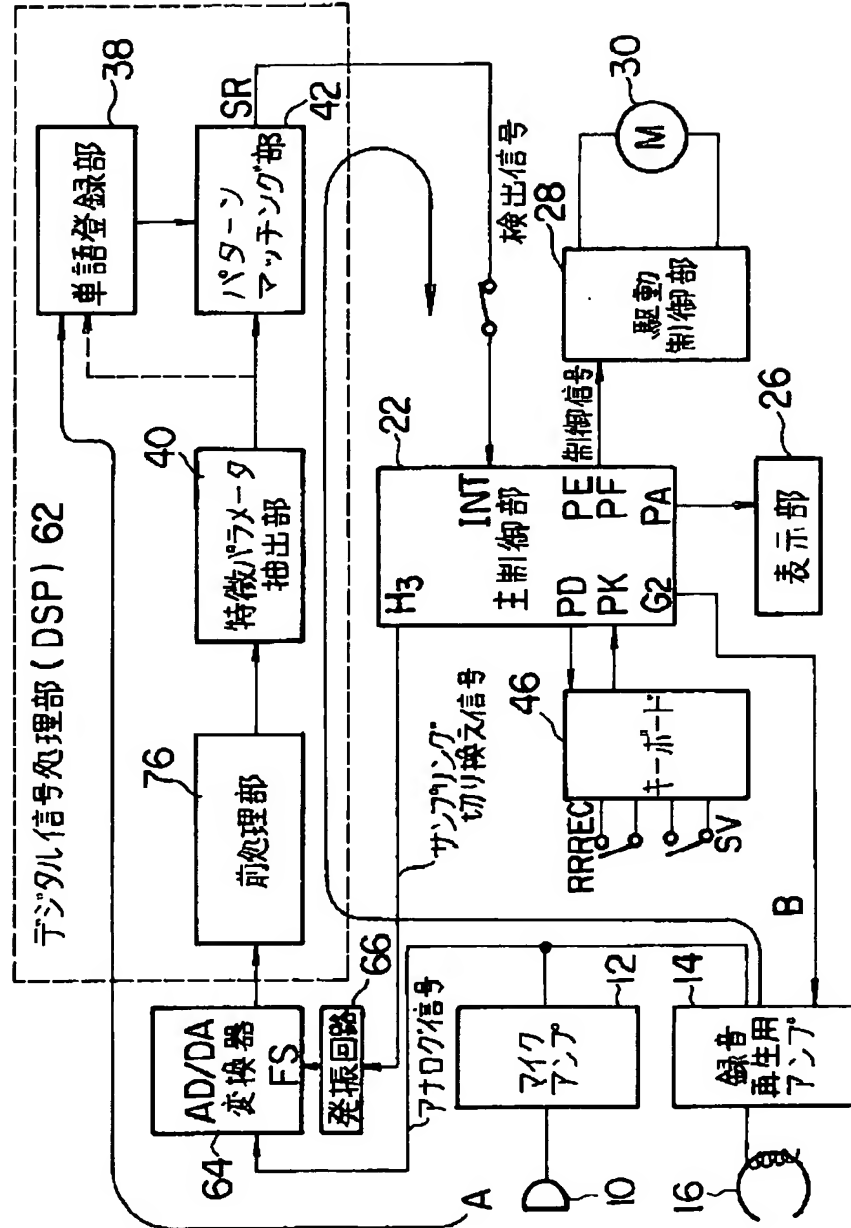




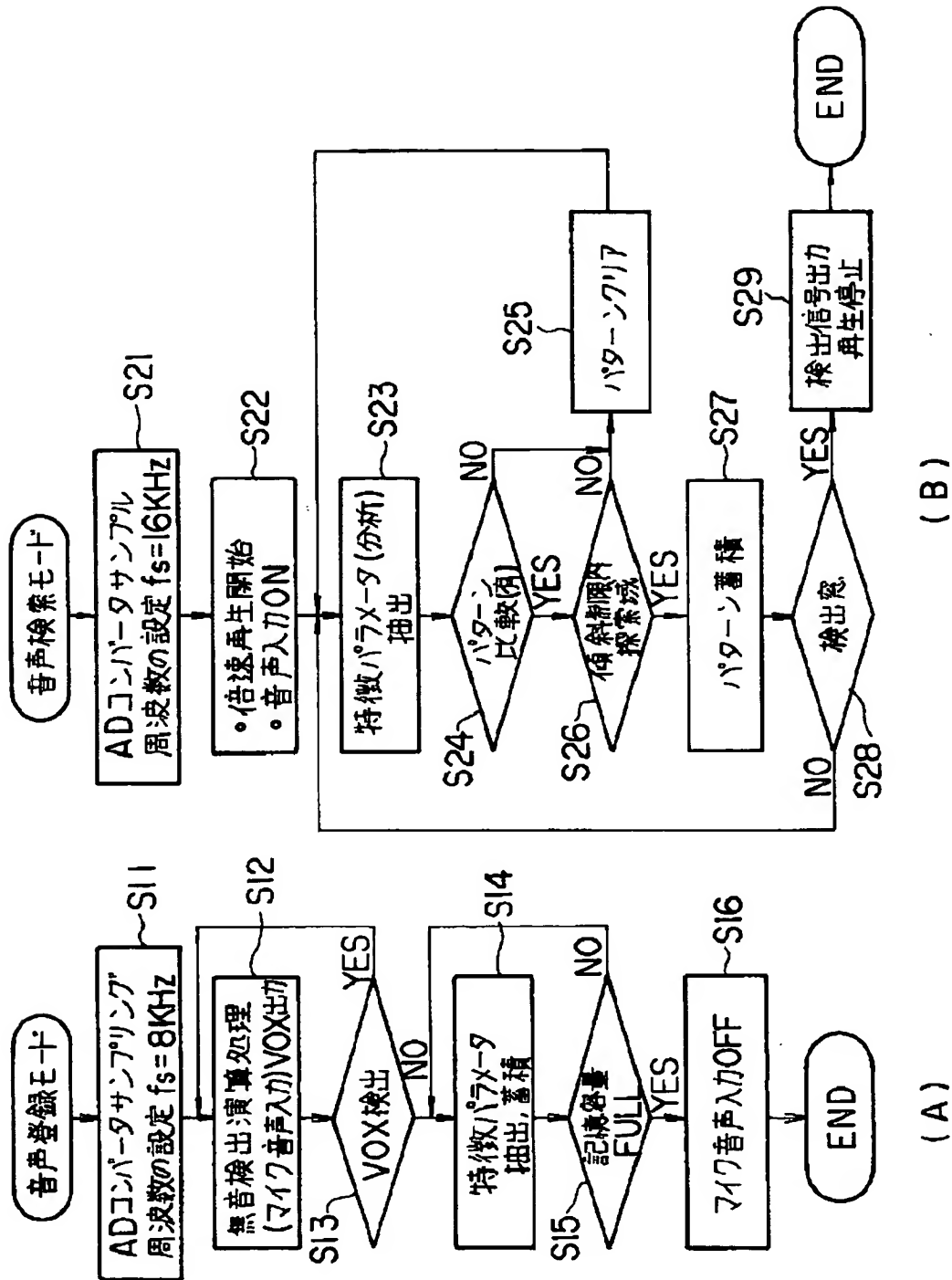
【図3】



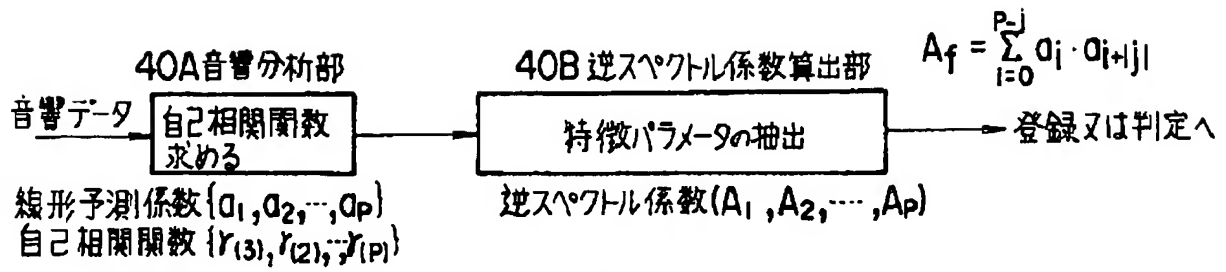
【図4】



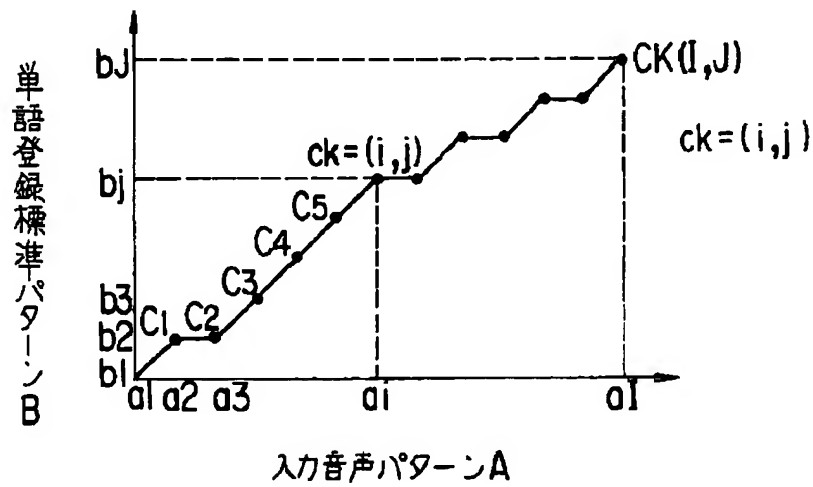
【図5】



【図6】

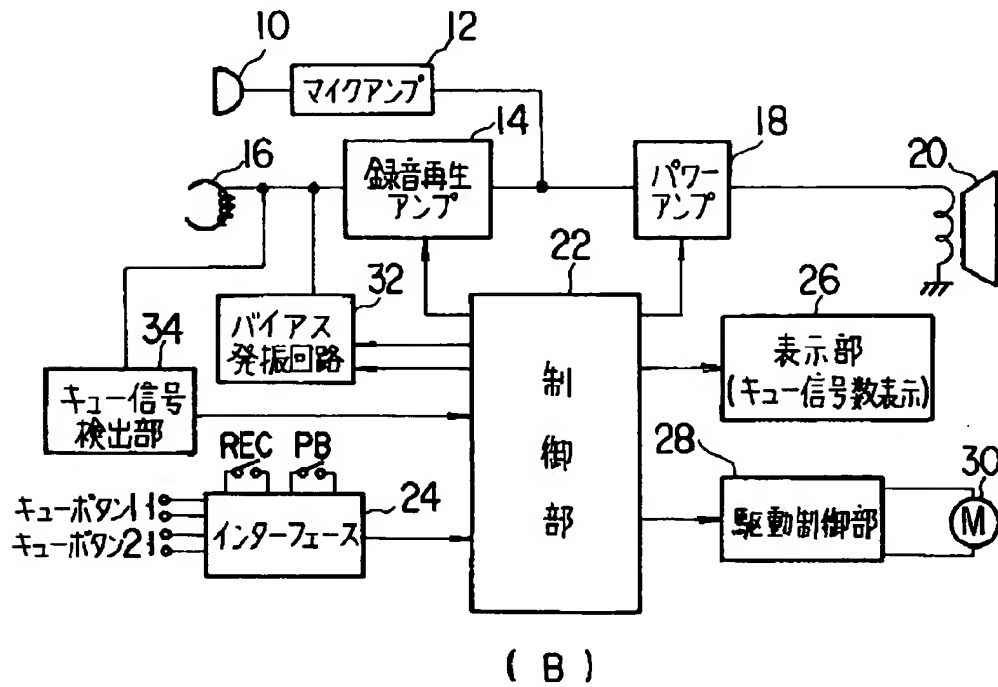
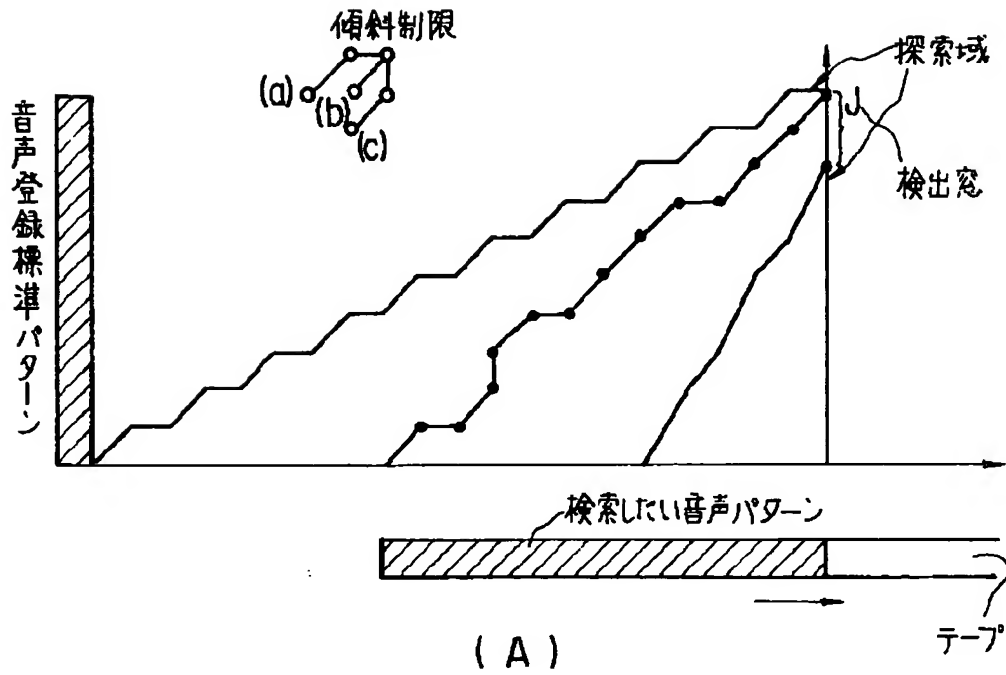


( A )

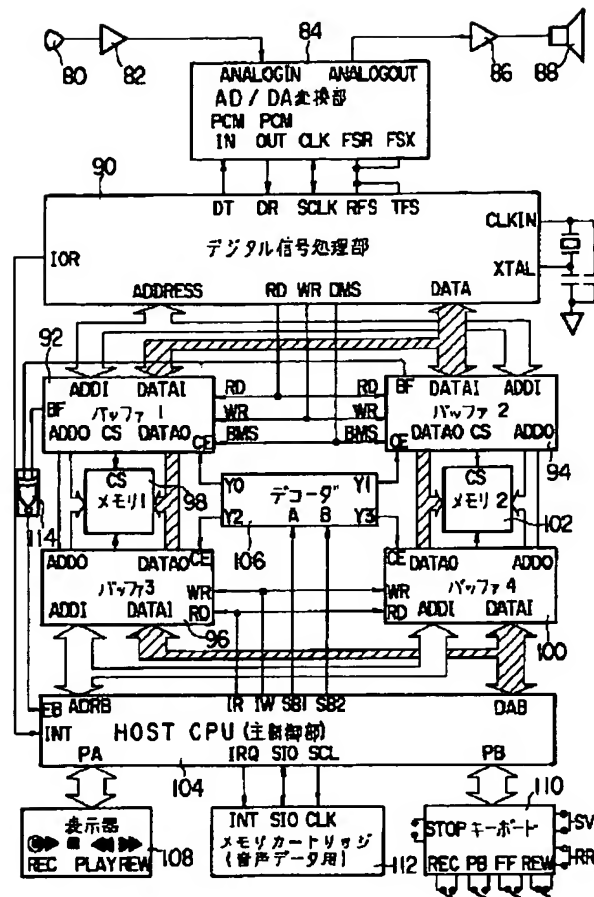


( B )

【図7】



【図8】



【図9】

